



**HAL**  
open science

# Exact calculation of false discovery rate for step-up and step-down procedures

Etienne Roquain, Fanny Villers

► **To cite this version:**

Etienne Roquain, Fanny Villers. Exact calculation of false discovery rate for step-up and step-down procedures. 2010. hal-00456385v1

**HAL Id: hal-00456385**

**<https://hal.science/hal-00456385v1>**

Preprint submitted on 15 Feb 2010 (v1), last revised 2 Apr 2010 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## EXACT CALCULATION OF FALSE DISCOVERY RATE FOR STEP-UP AND STEP-DOWN PROCEDURES

BY ETIENNE ROQUAIN AND FANNY VILLERS

We derive exact formulas for the false discovery rate (FDR) and the Power of any step-up and step-down procedures and for the  $s$ -th moment of the false discovery proportion (FDP) of any step-up procedure. These formulas are “explicit” in the sense that they only involve the parameters of the model and elementary operations. The  $p$ -values are assumed either independent or coming from an equicorrelated multivariate normal model and an additional mixture model for the true/false hypotheses is used. We propose several applications of these formulas. In particular, some new results related to least/most favorable configurations for the FDR and the variance of the FDP are presented.

**1. Introduction.** When testing simultaneously  $m$  null hypotheses, the false discovery proportion (FDP) is defined as the number of errors among all the rejected hypotheses and the false discovery rate (FDR) is defined as the average of the FDP. Since its introduction by [1], the FDR has become a widely used type I error criterion, because it is adaptive to the number of rejected hypotheses. However, the randomness of the denominator in the FDP expression makes the FDR somewhat mathematically difficult to study.

There is a considerable number of papers that deal with the FDR control under different dependency conditions between the test statistics (see for instance [1, 3–5, 22]). In the latter, the goal is, given a prespecified level  $\alpha$ , to provide a procedure with a FDR smaller than  $\alpha$  (for any value of the data law in a given distribution subspace, e.g. for some dependency assumptions). For instance, the famous linear step-up procedure (LSU), based on the Simes’s line (see [24]), has been proved to control the FDR under independence and under a positive dependence assumption (see [1, 3]). While controlling the FDR, one want to maximize the power of the procedure, the power being generally defined as the averaged number of erroneous rejections.

In this paper, we deal with the “reversed” approach: given the procedure, how to compute its FDR or its Power? For a constant thresholding and under a mixture model assuming independence between the test statistics, [25] addressed this question, while introducing the positive false discovery rate (pFDR). For step-up and step-down procedures, [11] computed the exact distribution of the rejection number when all the hypotheses are true and the test statistics are independent. Later, an asymptotic computation of the rejection number distribution of the LSU procedure (when  $m$  tends to infinity) has been advocated in [6] while introducing the criticality

---

*AMS 2000 subject classifications:* Primary 62J15; secondary 62G10; tertiary 60C05

*Keywords and phrases:* False Discovery Rate, multiple testing, step-up procedure, positive correlations, Power, equicorrelated multivariate normal distribution

phenomenon. Still considering the LSU procedure under the independence assumption, [9] gave a first exact expression for the moment of the FDP. In the context of positively correlated  $p$ -values following an equicorrelated multivariate normal model, [10] provided an asymptotic computation of the FDR of the LSU procedure as  $m$  grows to infinity, in a particular “Dirac-uniform” configuration for the  $p$ -values (that is when the  $p$ -values associated to false hypotheses are equal to 0). For other motivations, [21] and [15] also provided exact expressions for the FDR. Finally, recent works exhibit a connection between the FDR expression and the rejection number distribution in the step-up case (see, e.g. [4, 20, 22]).

The new contributions of the present paper are as follows:

- A new formula for the moment of the FDP of any step-up procedure is provided under independence. In particular, the first moment provides an expression for the FDR. Compared to the expression found in [9], our formula pushes the computation one step further, by using the exact distribution of the rejection number of any step-up procedure. Also, we put forward that the formula in [9] is wrong for moments of order greater than 3, and we provide a correction. We also give an expression for the power, which yields a considerably less complex calculation than in [14]. See Section 3.1.
- A new formula for the FDR and the power of any step-down procedure is presented under independence. To our knowledge, this expression is the first one that clearly relates the FDR to the number of rejections for a step-down procedure. We therefore obtain an explicit formula, which may be considered as a main contribution of the paper. See Section 3.1.
- Our formulas can be easily extended to the case where the  $p$ -value family follows an equicorrelated multivariate normal model, by a simple modification of the formulas valid under independence. See Section 3.2. However, this requires the use of a nonnegative correlation. The case of a possibly negative correlation is considered when only two hypotheses are tested. See Section 3.3.
- Our formulas are shown to corroborate the classical multiple testing results while they give rise to several new investigations, see Section 4. For instance, for the linear step-down procedure we prove that a  $p$ -value configuration maximizing the FDR, i.e. a least favorable configuration (LFC) for the FDR, is the “Dirac-uniform” configuration under independence and we provide an explicit upper-bound. Additionally, we found an exact expression of the minimum and the maximum of the variance of the FDP of the LSU over some  $p$ -value configurations of interest. The latter allows to better understand the asymptotic behavior of the FDP around the FDR when  $m$  tends to infinity.

All our formulas are valid nonasymptotically, that is, they hold for each  $m \geq 2$ . As a counterpart, they inevitably have a general form that can appear somewhat complex at first sight. Namely, denoting by  $\Psi_m$  the c.d.f. of the order statistics of  $m$  i.i.d. uniform variables on  $(0, 1)$ , the FDR formula for step-up procedures requires the computation of  $\Psi_m$  at a given point of  $\mathbb{R}^m$ , while the FDR formula for step-down procedures requires the computation of  $\Psi_m$  at  $m$  different points of  $\mathbb{R}^m$  (at most). However, let us underline what are to our opinion the two main interests

of this exact approach (see Section 4 for illustrations):

- For some model parameter configurations and after possible simplifications, the formulas are usable for further theoretical studies (monotonicity with respect to a parameter, convergence when  $m$  tends to infinity,...).
- For  $m$  not too large (say  $m \leq 1000$ ), these formulas can be fully computed, e.g. plotting exact graphs. Thus, they avoid using cumbersome and less accurate simulation experiments (extensively used in multiple testing literature).

## 2. Preliminaries.

2.1. *Models for the  $p$ -value family.* On a given probability space, we consider a finite set of  $m \geq 2$  null hypotheses, tested by a family of  $m$   $p$ -values  $\mathbf{p} = (p_i, i \in \{1, \dots, m\})$ . In this paper, for simplicity, we skip somewhat the formal definition of  $p$ -values by defining directly a  $p$ -value model, that is, by specifying the joint distribution of  $\mathbf{p}$ .

In what follows, we denote  $\mathcal{F}$  the set containing c.d.f.'s from  $[0, 1]$  into  $[0, 1]$  and  $F_0(t) = t$  the c.d.f. of the uniform distribution over  $[0, 1]$  (we chose to restrict our attention to the case where  $F_0(t) = t$  only for simplicity, see Section 5).

DEFINITION 2.1 (Conditional  $p$ -value models). • *The  $p$ -value family  $\mathbf{p}$  follows the conditional independent model with parameters  $H = (H_i)_{1 \leq i \leq m} \in \{0, 1\}^m$  and  $F_1 \in \mathcal{F}$ , that we denote by  $\mathbf{p} \sim P_{(H, F_1)}^I$ , if  $\mathbf{p} = (p_i, i \in \{1, \dots, m\})$  is a family of mutually independent variables and for all  $i$ ,*

$$p_i \sim \begin{cases} F_0 & \text{if } H_i = 0 \\ F_1 & \text{if } H_i = 1 \end{cases} .$$

- *The  $p$ -value family  $\mathbf{p}$  follows the conditional equicorrelated multivariate normal model (say for short, conditional EMN model) with parameters  $H = (H_i)_{1 \leq i \leq m} \in \{0, 1\}^m$ ,  $\rho \in [-(m-1)^{-1}, 1]$ , and  $\mu > 0$ , that we denote by  $\mathbf{p} \sim P_{(H, \rho, \mu)}^N$ , if for all  $i$ ,  $p_i \sim \bar{\Phi}(X_i + \mu H_i)$ , where the vector  $(X_i)_{1 \leq i \leq m}$  is distributed as a  $\mathbb{R}^m$ -valued Gaussian vector with zero means and a covariance matrix having 1 on the diagonal and  $\rho$  elsewhere and where  $\bar{\Phi}$  denotes the standard Gaussian distribution tail, that is,  $\bar{\Phi}(z) = \mathbb{P}[Z \geq z]$  for  $Z \sim \mathcal{N}(0, 1)$ . In that model, the marginal distributions of the  $p$ -values are thus given by:*

$$p_i \sim \begin{cases} F_0 & \text{if } H_i = 0 \\ F_1(t) = \bar{\Phi}(\bar{\Phi}^{-1}(t) - \mu) & \text{if } H_i = 1 \end{cases} .$$

The two above models are said “conditional” because the distribution of the  $p$ -values are defined conditionally on the value of the parameter  $H = (H_i)_{1 \leq i \leq m} \in \{0, 1\}^m$ . The latter codes for which hypotheses are true or false:  $H_i = 0$  (resp. 1) if the  $i$ -th null hypothesis is true (resp. false). We then denote by  $\mathcal{H}_0(H) := \{i \in \{1, \dots, m\} \mid H_i = 0\}$  the set corresponding

to the true null hypotheses and by  $m_0(H) := |\mathcal{H}_0(H)|$  its cardinal. Analogously, we define  $\mathcal{H}_1(H) := \{i \in \{1, \dots, m\} \mid H_i = 1\}$  and  $m_1(H) := |\mathcal{H}_1(H)| = m - m_0(H)$ .

To each one of the above models, we associated the “unconditional” version where we endow the parameter  $H$  with the prior distribution  $\mathcal{B}(1 - \pi_0)^{\otimes m}$ , making  $(H_i)_{1 \leq i \leq m} \in \{0, 1\}^m$  being a sequence of i.i.d. Bernoulli with parameter  $1 - \pi_0$ . On an intuitive point of view, this means that each hypothesis is true with probability  $\pi_0$ , independently from the other hypotheses. We thus define the following models for  $\mathbf{p}$  (or more precisely for  $(H, \mathbf{p})$ ):

- DEFINITION 2.2 (Unconditional  $p$ -value models). • *The couple  $(H, \mathbf{p})$  follows the unconditional independent model  $\overline{P}_{(\pi_0, F_1)}^I$  with parameters  $\pi_0 \in [0, 1]$  and  $F_1 \in \mathcal{F}$ , that we denote by  $(H, \mathbf{p}) \sim \overline{P}_{(\pi_0, F_1)}^I$  if  $H \sim \mathcal{B}(1 - \pi_0)^{\otimes m}$  and the distribution of  $\mathbf{p}$  conditionally to  $H$  is  $P_{(H, F_1)}^I$ , that is, conditionally on  $H$ ,  $\mathbf{p}$  follows the conditional independent model with parameters  $(H, F_1)$ . In that model, the  $p$ -values are i.i.d. with common c.d.f.  $G(t) = \pi_0 F_0(t) + (1 - \pi_0) F_1(t)$ .*
- *The couple  $(H, \mathbf{p})$  follows the unconditional equicorrelated multivariate normal model (say for short, unconditional EMN model) with parameters  $\pi_0 \in [0, 1]$ ,  $\rho \in [-(m - 1)^{-1}, 1]$ , and  $\mu > 0$ , that we denote by  $(H, \mathbf{p}) \sim \overline{P}_{(\pi_0, \rho, \mu)}^N$ , if  $H \sim \mathcal{B}(1 - \pi_0)^{\otimes m}$  and the distribution of  $\mathbf{p}$  conditionally on  $H$  is  $P_{(H, \rho, \mu)}^N$ , that is, conditionally on  $H$ ,  $\mathbf{p}$  follows the conditional EMN model with parameters  $(H, \rho, \mu)$ .*

An important point is that the quantities  $m_0(H)$  and  $m_1(H)$  are deterministic in the conditional models  $P^I, P^N$ , while they become random in the unconditional models  $\overline{P}^I, \overline{P}^N$  with  $m_0(H) \sim \mathcal{B}(m, \pi_0)$  and  $m_1(H) \sim \mathcal{B}(m, 1 - \pi_0)$ .

The conditional independent model is one of the most standard  $p$ -value model and was for instance considered in the original paper of [1]. Its unconditional version, also called the “random effects model”, is very convenient and has been widely used since its introduction by [8], see for instance [13, 26].

The conditional EMN model is one of the most simple instances of model introducing dependencies between the  $p$ -values. It corresponds to a one-sided testing on the mean of  $X_i + \mu H_i$ , simultaneously for all  $1 \leq i \leq m$ . It has become quite standard in recent FDR multiple testing literature; for instance, it was used in [10] with  $\mu = \infty$  and it has been considered in [2, 5] for numerical experiments. In this model, provided that  $\rho \geq 0$ , the  $p$ -values are positively regression dependent on each one on the subset  $\mathcal{H}_0(H)$  (PRDS on  $\mathcal{H}_0(H)$ ) which is one dependency condition that suffices for FDR control (see [3]). The unconditional version of this model is convenient because it provides exchangeable  $p$ -values (although not independent when  $\rho \neq 0$ ).

Additionally, we will sometimes consider the “Dirac-uniform configuration” for the above models. In that configuration, all the  $p$ -values corresponding to false nulls ( $H_i = 1$ ) are equal to zero, that is,  $F_1$  is constantly equal to 1 for the independent models and  $\mu = \infty$  for the EMN models. This configuration was introduced in [10] to increase the FDR as much as possible for

the linear step-up procedures and thus appears as a “least favorable configuration” for the FDR (see also Theorem 5.3 in [3]).

*2.2. Multiple testing procedures, FDR and power.* A *multiple testing procedure*  $R$  is defined as an algorithm which, from the data, aims to reject part of the null hypotheses. Below, we will consider, as is usually the case, multiple testing procedures which can be written as a function of the  $p$ -value family  $\mathbf{p} = (p_i, i \in \{1, \dots, m\})$ . More formally, we define a multiple testing procedure as a measurable function  $R$ , which takes as input a realization of the  $p$ -value family  $\mathbf{p} \in [0, 1]^m$  and which returns a subset  $R(\mathbf{p})$  of  $\{1, \dots, m\}$ , corresponding to the rejected hypotheses (i.e.  $i \in R(\mathbf{p})$  means that the  $i$ -th hypothesis is rejected by  $R$  for the observed  $p$ -values  $\mathbf{p}$ ).

Particular multiple testing procedures are step-up and step-down procedures. First define a *threshold* as any nondecreasing sequence  $\mathbf{t} = (t_k)_{1 \leq k \leq m} \in [0, 1]^m$  (with  $t_0 = 0$  by convention). Next, for any threshold  $\mathbf{t}$ , the *step-up procedure* of threshold  $\mathbf{t}$ , denoted here by  $SU(\mathbf{t})$ , rejects the  $i$ -th hypothesis if  $p_i \leq t_{\hat{k}}$ , with  $\hat{k} = \max\{k \in \{0, 1, \dots, m\} \mid p_{(k)} \leq t_k\}$ , where  $p_{(1)} \leq p_{(2)} \leq \dots \leq p_{(m)}$  denote the ordered  $p$ -values (with the convention  $p_{(0)} = 0$ ). In particular, the procedure  $SU(\mathbf{t})$  using  $t_k = \alpha k/m$  corresponds to the standard linear step-up procedure of [1], denoted here by LSU. A less rejecting procedure uses a step-down algorithm; for any threshold  $\mathbf{t}$ , the *step-down procedure* of threshold  $\mathbf{t}$ , denoted here by  $SD(\mathbf{t})$ , rejects the  $i$ -th hypothesis if  $p_i \leq t_{\tilde{k}}$ , with  $\tilde{k} = \max\{k \in \{0, 1, \dots, m\} \mid \forall k' \leq k, p_{(k')} \leq t_{k'}\}$ . Analogously to the step-up case, the procedure  $SD(\mathbf{t})$  using  $t_k = \alpha k/m$  is called the linear step-down procedure and is denoted here by LSD.

Next, associated to any multiple testing procedure  $R$  and any configuration of true/false hypotheses  $H \in \{0, 1\}^m$ , we introduce the false discovery proportion (FDP) of  $R$  as the proportion of true hypotheses in the set of the rejected hypotheses, that is,

$$(1) \quad \text{FDP}(R, H) = \frac{|\mathcal{H}_0(H) \cap R|}{|R| \vee 1},$$

where  $|\cdot|$  denotes the cardinality function. Then, while for any multiple testing procedure  $R$ , the false discovery rate (FDR) is defined as the mean of the FDP (see [1]), the power is (generally) defined as the expected number of correctly rejected hypotheses divided by the number of false hypotheses. Of course, the FDR and the power depend on the distribution that generates the  $p$ -values, and we may use the models defined in Section 2.1. Formally, for any distribution  $P$  coming from a conditional model using parameter  $H \in \{0, 1\}^m$ , we let

$$(2) \quad \text{FDR}(R, P) = \mathbb{E}_{\mathbf{p} \sim P}[\text{FDP}(R(\mathbf{p}), H)],$$

$$(3) \quad \text{Pow}(R, P) = m_1(H)^{-1} \mathbb{E}_{\mathbf{p} \sim P}[|\mathcal{H}_1(H) \cap R(\mathbf{p})|].$$

Similarly, for any  $p$ -value distribution  $\overline{P}$  coming from an unconditional model, the FDR and the Power use an additional averaging over  $H \sim \mathcal{B}(1 - \pi_0)^{\otimes m}$  and are defined by:

$$(4) \quad \text{FDR}(R, \overline{P}) = \mathbb{E}_{(H, \mathbf{p}) \sim \overline{P}}[\text{FDP}(R(\mathbf{p}), H)],$$

$$(5) \quad \text{Pow}(R, \overline{P}) = (\pi_1 m)^{-1} \mathbb{E}_{(H, \mathbf{p}) \sim \overline{P}}[|\mathcal{H}_1(H) \cap R(\mathbf{p})|].$$

Remark that, for convenience, (5) is not exactly defined as the expectation of (3), because of the denominator. It corresponds precisely to the expected number of correctly rejected hypotheses divided by the *expected* number of false hypotheses.

In the paper, to simplify the notation, we sometimes drop the explicit dependency in  $\mathbf{p}$ ,  $H$  or  $P$ , writing e.g.  $R$  instead of  $R(\mathbf{p})$ ,  $\mathcal{H}_0$  instead of  $\mathcal{H}_0(H)$  and  $\text{FDR}(R)$  instead of  $\text{FDR}(R, P)$ .

**2.3. Some notation and useful results.** For any  $k \geq 0$  and any threshold  $\mathbf{t} = (t_1, \dots, t_k)$ , we denote

$$(6) \quad \Psi_k(\mathbf{t}) = \Psi_k(t_1, \dots, t_k) = \mathbb{P} [U_{(1)} \leq t_1, \dots, U_{(k)} \leq t_k].$$

where  $(U_i)_{1 \leq i \leq k}$  is a sequence of  $k$  variables i.i.d. uniform on  $(0, 1)$  and with the convention  $\Psi_0(\cdot) = 1$ . In practice, quantity (6) can be evaluated using Bolshev's recursion  $\Psi_k(\mathbf{t}) = 1 - \sum_{i=1}^k \binom{k}{i} (1 - t_{k-i+1})^i \Psi_{k-i}(t_1, \dots, t_{k-i})$  or Steck's recursion  $\Psi_k(\mathbf{t}) = (t_k)^k - \sum_{j=0}^{k-2} \binom{k}{j} (t_k - t_{j+1})^{k-j} \Psi_j(t_1, \dots, t_j)$  (see [23], p. 366-369). Additionally, the following relation holds (see Lemma 2.1 in [11]): for all  $k \in \mathbb{N}$  and  $\nu_1, \nu_2 \in \mathbb{R}$  such that  $0 \leq \nu_1 + \nu_2 \leq \nu_1 + k\nu_2 \leq 1$ ,

$$(7) \quad \Psi_k(\nu_1 + \nu_2, \dots, \nu_1 + k\nu_2) = (\nu_1 + \nu_2)(\nu_1 + (k+1)\nu_2)^{k-1}.$$

From the  $\Psi_k$ , we define the following useful quantities: for any threshold  $\mathbf{t} = (t_k)_{1 \leq k \leq m}$  and  $k \geq 0$ ,  $k \leq m$ , we let

$$(8) \quad \mathcal{D}_m(\mathbf{t}, k) = \binom{m}{k} (t_k)^k \Psi_{m-k}(1 - t_m, \dots, 1 - t_{k+1}),$$

$$(9) \quad \tilde{\mathcal{D}}_m(\mathbf{t}, k) = \binom{m}{k} (1 - t_{k+1})^{m-k} \Psi_k(t_1, \dots, t_k).$$

Above, note that  $(t_k)^k$  and  $(1 - t_{k+1})^{m-k}$  are correct even if  $(t_j)_j$  is only defined for  $1 \leq j \leq m$ . Note that Bolshev's recursion provides  $\sum_{k=0}^m \mathcal{D}_m(\mathbf{t}, k) = \sum_{k=0}^m \tilde{\mathcal{D}}_m(\mathbf{t}, k) = 1$  for any threshold  $\mathbf{t}$ .

Next, define for any threshold  $\mathbf{t} = (t_j)_{1 \leq j \leq m+1}$ , and  $k, k' \geq 1$ ,  $k \leq k' \leq m$ ,

$$(10) \quad \begin{aligned} \tilde{\mathcal{D}}_m(\mathbf{t}, k, k') &= \tilde{\mathcal{D}}_m((t_j)_{1 \leq j \leq m}, k) \tilde{\mathcal{D}}_{m-k} \left( \left( \frac{t_{k+1+j} - t_{k+1}}{1 - t_{k+1}} \right)_{1 \leq j \leq m-k}, k' - k \right) \\ &= \binom{m}{k} \binom{m-k}{k' - k} (1 - t_{k'+2})^{m-k'} \Psi_k(t_1, \dots, t_k) \Psi_{k'-k}(t_{k+2} - t_{k+1}, \dots, t_{k'+1} - t_{k+1}). \end{aligned}$$

We easily check that  $\sum_{k=1}^m \sum_{k'=k}^m \tilde{\mathcal{D}}_m(\mathbf{t}, k, k') = 1$  for any threshold  $\mathbf{t}$ .

Finally, we will use the so-called Stirling numbers of the second kind, defined as coefficients  $\{s\}_\ell$  for  $s, \ell \geq 1$  by  $\{s\}_0 = 0$ ,  $\{s\}_\ell = 0$  for  $\ell > s$ ,  $\{1\}_1 = 1$  and the recurrence relation: for all

$$1 \leq \ell \leq s + 1,$$

$$(11) \quad \left\{ \begin{matrix} s+1 \\ \ell \end{matrix} \right\} = \ell \left\{ \begin{matrix} s \\ \ell \end{matrix} \right\} + \left\{ \begin{matrix} s \\ \ell-1 \end{matrix} \right\}.$$

Some values of  $\left\{ \begin{matrix} s \\ \ell \end{matrix} \right\}$  are given in Table 1. The coefficient  $\left\{ \begin{matrix} s \\ \ell \end{matrix} \right\}$  counts the number of ways to partition a set of  $s$  elements into  $\ell$  nonempty subsets. For instance, they are useful to compute the  $s$ -th moment of a binomial distribution: if  $m_0 \sim \mathcal{B}(m, \pi_0)$ , we get that  $\mathbb{E}[m_0^s] = \sum_{\ell=1}^{s \wedge m} \frac{m!}{(m-\ell)!} \left\{ \begin{matrix} s \\ \ell \end{matrix} \right\} \pi_0^\ell$  (this easily derives from Lemma A.3 in Appendix A).

$s \setminus \ell$	0	1	2	3	4	5	6
1	0	1	0	0	0	0	0
2	0	1	1	0	0	0	0
3	0	1	3	1	0	0	0
4	0	1	7	6	1	0	0
5	0	1	15	25	10	1	0

TABLE 1  
Values for  $\left\{ \begin{matrix} s \\ \ell \end{matrix} \right\}$ ,  $0 \leq \ell \leq s + 1 \leq 6$ .

**3. Main results.** In this section, we provide our main results, that is, exact expressions for the FDR (and more generally for the  $s$ -th moment of the FDP in the step-up case) and the power, for any step-up and step-down procedures.

3.1. *Unconditional independent model,  $m \geq 2$ .* We first examine the step-up case. On the one hand, the authors of [9] have expressed the  $s$ -th moment of the FDP of the LSU in function of the distribution of  $|\text{LSU}|$  (i.e. the rejection number of LSU), or more precisely, after a slight modifications of it (actually, their formula needs additional terms for  $s \geq 3$ , see Remark 3.2 below). However, the exact distribution of  $|\text{LSU}|$  is unknown in the conditional independent model and their expression is not fully computable. On the other hand, [11] derived the exact distribution of the rejection number of any step-up procedure in the unconditional model, using that the  $p$ -values are in that case *i.i.d.* In our framework, the latter can be expressed as follows: denoting  $G(t) = \pi_0 F_0(t) + (1 - \pi_0) F_1(t)$  the common c.d.f. of the  $p$ -values (in the unconditional independent model), we have for  $0 \leq k \leq m$  that

$$(12) \quad \mathbb{P}[|\text{SU}(\mathbf{t})| = k] = \mathcal{D}_m([G(t_j)]_{1 \leq j \leq m}, k).$$

(actually, their result stands for  $\pi_0 = 1$ , but the generalization is straightforward).

Our first main result adapts the result of [9] to the unconditional model (and for a general  $\text{SU}(\mathbf{t})$ , not only for LSU) and combines it with (12), to obtain a fully computable expression.

**THEOREM 3.1.** *When testing  $m \geq 2$  hypotheses, consider a step-up procedure  $SU(\mathbf{t})$  with threshold  $\mathbf{t}$  and the notation of Section 2.3. Then for any parameters  $\pi_0 \in [0, 1]$  and  $F_1 \in \mathcal{F}$ , denoting  $G(t) = \pi_0 F_0(t) + \pi_1 F_1(t)$ , we have for any  $1 \leq s \leq m$ , under the generating distribution  $(H, \mathbf{p}) \sim \overline{P}_{(\pi_0, F)}^I$  of the unconditional independent model, denoting by  $\{s_\ell\}$  the Stirling number of second kind,*

$$(13) \quad \mathbb{E}[FDP(SU(\mathbf{t}), H)^s] = \sum_{\ell=1}^s \frac{m!}{(m-\ell)!} \{s_\ell\} \pi_0^\ell \sum_{k=\ell}^m \frac{F_0(t_k)^\ell}{k^s} \mathcal{D}_{m-\ell}([G(t_{j+\ell})]_{1 \leq j \leq m-\ell}, k-\ell)$$

$$(14) \quad \text{Pow}(SU(\mathbf{t}), \overline{P}_{(\pi_0, F_1)}^I) = \sum_{k=1}^m F_1(t_k) \mathcal{D}_{m-1}([G(t_{j+1})]_{1 \leq j \leq m-1}, k-1).$$

In particular, we have

$$(15) \quad \text{FDR}(SU(\mathbf{t}), \overline{P}_{(\pi_0, F_1)}^I) = \pi_0 m \sum_{k=1}^m \frac{F_0(t_k)}{k} \mathcal{D}_{m-1}([G(t_{j+1})]_{1 \leq j \leq m-1}, k-1)$$

$$(16) \quad = \pi_0 \sum_{k=1}^m \binom{m}{k} F_0(t_k) (G(t_k))^{k-1} \Psi_{m-k}(1 - G(t_m), \dots, 1 - G(t_{k+1})).$$

To put some rationale behind the above formulas, let us choose some particular parameter configurations: first, taking  $F_1 = 1$  in the RHS of (14) gives  $\text{Pow}(SU(\mathbf{t}), \overline{P}_{(\pi_0, F_1)}^I) = 1$  which is obviously true from the definition of the power. Second, by taking  $t_k = \alpha k/m$ , (15) reduces to the well known result of [3]: the FDR of the linear step-up procedure is equal to  $\pi_0 \alpha$ , that is,  $\text{FDR}(\text{LSU}, \overline{P}_{(\pi_0, F_1)}^I) = \pi_0 \alpha$  (actually, this result was initially established in the conditional model). Third, considering a threshold  $\mathbf{t}(\cdot)$  constantly equal to  $t \in [0, 1]$  into (16), we obtain that  $\text{FDR}(t, \overline{P}_{\pi_0, F}^I) = \pi_0 t \sum_{k=1}^m \binom{m}{k} (G(t))^{k-1} (1 - G(t))^{m-k} = \pi_0 \frac{t}{G(t)} (1 - (1 - G(t))^m)$ , which recovers the well known result of [26] (Theorem 1 therein). Also note that for this constant threshold  $t$ , the power (14) reduces to  $F_1(t)$ , which is in accordance with the definition of the power. Fourth, taking  $\pi_0 = 1$  in the LHS of (13) gives  $P(|\text{SU}(\mathbf{t})| > 0)$  because in that case  $\text{FDP} \in \{0, 1\}$ . We can check that the RHS of (13) can be written as  $\sum_{k=1}^m t_k^k \Psi_{m-k}(1 - t_m, \dots, 1 - t_{k+1}) \sum_{\ell=1}^k k^{-s} \frac{m!}{(m-\ell)!} \{s_\ell\} \binom{m-\ell}{k-\ell}$  and we coherently obtain  $\sum_{k=1}^m \mathbb{P}[|\text{SU}(\mathbf{t})| = k]$  by using the formula  $k^s = \sum_{\ell=1}^{s \wedge k} \frac{k!}{(k-\ell)!} \{s_\ell\}$ .

Next, (14) provides an exact expression of the power of a step-up procedure in the unconditional model:  $\text{Pow}(\text{LSU}, \overline{P}_{(\pi_0, F_1)}^I) = \sum_{k=1}^m F_1(t_k) \binom{m-1}{k-1} (G(t_k))^{k-1} \Psi_{m-k}(1 - G(t_m), \dots, 1 - G(t_{k+1}))$ . The author [14] have obtained an exact expression of the power of the LSU ( $t_k = \alpha k/m$ ) under independence (in the conditional model), but their formula was reported to have a complexity exponential in  $m$ , which is intractable for large  $m$ . Here, we obtained a much less complex formula, requiring (at most) the computation of the function  $\Psi_m$  at one point of  $\mathbb{R}^m$ .

Several other applications of Theorem 3.1 are presented in Section 3.2 and in Section 4.

REMARK 3.2. Compared to expression (2.1) in Theorem 2.1 of [9], formula (13) uses additional factors  $\{\ell^s\}$ . They are definitely needed as soon as  $s \geq 3$ ; for instance, taking  $t_m = 1$ , the corresponding step-up procedure rejects all the hypotheses and (13) reduces to the computation of the  $s$ -th moment of a binomial distribution, which uses at least one  $\{\ell^s\} > 1$ .

Now, let us consider the step-down case. An exact computation of the distribution of  $|\text{SD}(\mathbf{t})|$ , the rejection number of a step-down procedure of threshold  $\mathbf{t}$ , is done in formula (4) p.344 of [23] (see also [11]) in the unconditional independent model (initially for  $G$  equals to identity but the generalization is straightforward). The latter can be written as follows: for all  $k \in \{0, \dots, m\}$ ,

$$(17) \quad \mathbb{P}[|\text{SD}(\mathbf{t})| = k] = \tilde{\mathcal{D}}_m([G(t_j)]_{1 \leq j \leq m}, k).$$

However, to our knowledge, there exists no exact formula for step-down procedures making a general link between the FDR and distributions of the type (17). We show here that something more is required, namely, the joint distribution of  $(|\text{SD}(\mathbf{t})|, |\text{SD}(\mathbf{t}')|)$  where for  $1 \leq k \leq m$ ,  $t'_k := t_{k+1}$  (for any given threshold  $\mathbf{t}$  defined over  $\{1, \dots, m+1\}$  and given that the procedures  $|\text{SD}(\mathbf{t})|$  and  $|\text{SD}(\mathbf{t}')|$  are testing  $m$  null hypotheses). In Section 6.4, we proved the following: for any  $k, k'$  with  $0 \leq k \leq m$  and  $k \leq k' \leq m$ , we have

$$(18) \quad \mathbb{P}[|\text{SD}(\mathbf{t})| = k, |\text{SD}(\mathbf{t}')| = k'] = \tilde{\mathcal{D}}_m([G(t_j)]_{1 \leq j \leq m+1}, k, k').$$

We then obtain the following result:

THEOREM 3.3. For  $m \geq 2$  hypotheses, consider the unconditional independent model  $\bar{P}_{(\pi_0, F_1)}^I$ , a step-down procedure  $\text{SD}(\mathbf{t})$  with threshold  $\mathbf{t}$  and the notation of Section 2.3. Then for any parameters  $\pi_0 \in [0, 1]$  and  $F_1 \in \mathcal{F}$  (denoting  $G(t) = \pi_0 F_0(t) + \pi_1 F_1(t)$ ), we have

$$(19) \quad \text{FDR}(\text{SD}(\mathbf{t}), \bar{P}_{(\pi_0, F_1)}^I) = \pi_0 m \sum_{k=1}^m \sum_{k'=k}^m \frac{F_0(t_k)}{k'} \tilde{\mathcal{D}}_{m-1}([G(t_j)]_{1 \leq j \leq m}, k-1, k'-1),$$

$$(20) \quad \text{Pow}(\text{SD}(\mathbf{t}), \bar{P}_{(\pi_0, F_1)}^I) = \sum_{k=1}^m F_1(t_k) \tilde{\mathcal{D}}_{m-1}([G(t_j)]_{1 \leq j \leq m-1}, k-1).$$

To the best of our knowledge, (19) is the first exact expression for a step-down procedure that relies the FDR to the (joint) distribution of rejection numbers. The main tool to get this result is Lemma A.2.

One straightforward consequence of (19) is the upper-bound  $\pi_0 m \sum_{k=1}^m \frac{t_k}{k} \tilde{\mathcal{D}}_{m-1}([G(t_j)]_{1 \leq j \leq m}, k-1)$  for the FDR of  $\text{SD}(\mathbf{t})$ , which in particular proves that the FDR of  $\text{SD}(\mathbf{t})$  is always smaller than the one of  $\text{SU}(\mathbf{t})$  (as soon as  $t_k/k$  is nondecreasing in  $k$ ). While the latter result should probably be considered as well known, it is not trivial because when increasing the rejection number, both numerator and denominator are increasing within the FDP expression.

To illustrate ones more (19), we may consider the case where  $t_k = \alpha k/m$ , that is, we may compute exactly the FDR of the LSD procedure. In the Dirac-uniform model and using (7), elementary computations shows that the latter has the following expression:

$$(21) \quad \begin{aligned} \text{FDR}(\text{LSD}, \bar{P}_{(\pi_0, F_1=1)}^I) &= \pi_0 \alpha \left( \pi_1 + \pi_0 \frac{\alpha}{m} \right) \sum_{k=1}^m \sum_{j=k}^m \frac{k}{j} \binom{m-1}{k-1} \binom{m-k}{j-k} \pi_0^{m-k} \\ &\times \left( \pi_1 + \pi_0 \frac{\alpha k}{m} \right)^{k-2} \left( \frac{\alpha(j-k+1)}{m} \right)^{j-k-1} \left( 1 - \frac{\alpha(j+1)}{m} \right)^{m-j}. \end{aligned}$$

Some other applications of Theorem 3.3 are presented in Section 3.2 and in Section 4.

**3.2. Unconditional EMN model with  $\rho \in [0, 1]$ .** In this subsection, our goal is to obtain results similar to Theorem 3.1 and Theorem 3.3, but this time in the unconditional EMN model with a non negative correlation  $\rho \in [0, 1]$ . In that case, we easily see that, in the unconditional EMN model of parameters  $(\pi_0, \rho, \mu)$ , the joint distribution of the  $p$ -values can be realized as follows: for all  $i$ ,  $p_i = \bar{\Phi}(\sqrt{\rho} \bar{\Phi}^{-1}(U) + \sqrt{1-\rho} \bar{\Phi}^{-1}(U_i) + \mu H_i)$ , where  $U, (U_i)_i$  are all i.i.d. uniform on  $(0, 1)$  (and independent of  $(H_i)_i$ ) (this idea can be traced back to [18, 27]). As a consequence, conditionally on  $U = u$ , the  $p$ -values follows the unconditional independent model of parameters  $\pi_0, F_0(\cdot | u, \rho)$  and  $F_1(\cdot | u, \rho)$  where we let

$$(22) \quad F_0(t | u, \rho) = \bar{\Phi} \left( \frac{\bar{\Phi}^{-1}(t) - \sqrt{\rho} \bar{\Phi}^{-1}(u)}{\sqrt{1-\rho}} \right); \quad F_1(t | u, \rho) = \bar{\Phi} \left( \frac{\bar{\Phi}^{-1}(t) - \sqrt{\rho} \bar{\Phi}^{-1}(u) - \mu}{\sqrt{1-\rho}} \right).$$

for  $\rho \in [0, 1)$  and  $F_0(t | u, 1) = \mathbf{1}\{u \leq t\}$ ;  $F_1(t | u, 1) = \mathbf{1}\{u \leq \bar{\Phi}(\bar{\Phi}^{-1}(t) - \mu)\}$  for  $\rho = 1$ . This means that, to derive a formula in the unconditional EMN model, we may directly use formulas holding in the unconditional independent model (with the above modified c.d.f.'s) and using an additional integration over  $u \in (0, 1)$ . To illustrate this method, let us first consider the case of a constant thresholding  $\mathbf{t}(\cdot) = t \in [0, 1]$ . In that case, we obtained in the previous subsection that  $\text{FDR}(t, \bar{P}_{\pi_0, F_1}^I) = \pi_0 \frac{t}{G(t)} (1 - (1 - G(t))^m)$ . Hence, the latter reasoning yields, in the unconditional EMN model,  $\text{FDR}(t, \bar{P}_{(\pi_0, \rho, \mu)}^N) = \pi_0 t \int_0^1 \frac{1 - (1 - G(t | u, \rho))^m}{G(t | u, \rho)} du$ , where  $G(t | u, \rho) = \pi_0 F_0(t | u, \rho) + \pi_1 F_1(t | u, \rho)$  is defined using (22). More generally, we can deduce from Theorem 3.1 and Theorem 3.3 the following result (the formulas are not fully written for short):

**COROLLARY 3.4.** *For  $m \geq 2$  hypotheses, consider the unconditional EMN model  $\bar{P}_{(\pi_0, \rho, \mu)}^N$  with parameters  $\pi_0 \in [0, 1]$ ,  $\mu > 0$  and  $\rho \in [0, 1]$  and let  $G(t | u, \rho) = \pi_0 F_0(t | u, \rho) + \pi_1 F_1(t | u, \rho)$  using notation (22). Then, for any threshold  $\mathbf{t}$ , under the generating distribution  $(H, \mathbf{p}) \sim \bar{P}_{(\pi_0, \rho, \mu)}^N$ ,  $\mathbb{E}[\text{FDP}(SU(\mathbf{t}), H)^s]$  (resp.  $\text{Pow}(SU(\mathbf{t}))$ ;  $\text{FDR}(SU(\mathbf{t}))$ ) is given by the RHS of (13) (resp. (14); (15)), by replacing  $F_0(\cdot)$  by  $F_0(\cdot | u, \rho)$ ,  $F_1(\cdot)$  by  $F_1(\cdot | u, \rho)$  and  $G(\cdot)$  by  $G(\cdot | u, \rho)$ , and by integrating over  $u$  with respect to the Lebesgue measure on  $(0, 1)$ . Additionally, a similar result holds for step-down procedures using (19) and (20).*

In particular, applying Corollary 3.4 for the LSU procedure, we obtain that  $\text{FDR}(\text{LSU}, \overline{P}_{(\pi_0, \rho, \mu)}^N)$  equals:

$$(23) \quad \pi_0 \sum_{k=1}^m \binom{m}{k} \int_0^1 F_0(\alpha k/m | u, \rho) G(\alpha k/m | u, \rho)^{k-1} \Psi_{m-k}((1 - G(\alpha(m-j+1)/m | u, \rho))_{1 \leq j \leq m-k}) du.$$

An expression for  $\lim_m \text{FDR}(\text{LSU}, \overline{P}_{(\pi_0, \rho, \mu)}^N)$  was provided in [10], by considering the asymptotic framework where  $m$  tends to infinity. We compared the latter to our formula (23) by plotting the graph corresponding to their Figure 3 (not reported here). The results are qualitatively the same for  $\pi_0 < 1$ , but present major differences when  $\pi_0 = 1$  and  $\rho$  is small. This is in accordance with the simulations reported in Section 5 of [10]. Hence, the asymptotic analysis may not reflect what happens for a realistically finite  $m$ , which can be seen as a limitation with respect to our non-asymptotic approach. Furthermore, when  $\pi_0 = 1$ , [10] proved that  $\lim_{\rho \rightarrow 0} \lim_m \text{FDR}(\text{LSU}) = \overline{\Phi}(\sqrt{-2 \log \alpha}) < \alpha$  and they put forward the hypothesis that  $\lim_m \lim_{\rho \rightarrow 0} \text{FDR}(\text{LSU}) = \alpha$  (using simulations). They thus concluded that the order of limits plays a crucial role. The next result proves in particular that this hypothesis is true by showing that for each  $m$ ,  $\text{FDR}(\text{LSU})$  is continuous in  $\rho$ .

**COROLLARY 3.5.** *For any  $m \geq 2$  and for any threshold  $\mathbf{t}$ , the quantities  $\text{FDR}(\text{SU}(\mathbf{t}), \overline{P}_{(\pi_0, \rho, \mu)}^N)$  and  $\text{FDR}(\text{SD}(\mathbf{t}), \overline{P}_{(\pi_0, \rho, \mu)}^N)$  are continuous in any  $\pi_0 \in [0, 1]$ , any  $\rho \in [0, 1]$  and any  $\mu > 0$ .*

Corollary 3.5 is a straightforward consequence of Corollary 3.4; to prove the continuity in  $\rho = 1$ , we may remark that for any  $u$  outside the set  $\mathcal{S} = \{t_k, k \geq 1\} \cup \{\overline{\Phi}(\overline{\Phi}^{-1}(t_k) - \mu), k \geq 1\}$  of zero Lebesgue measure, the functions  $F_0(t | u, \rho)$  and  $F_1(t | u, \rho)$  are continuous in  $\rho = 1$ .

In particular, Corollary 3.5 shows that the limit of the FDR when  $\rho \rightarrow 1$  is given by the degenerated case  $\rho = 1$ . In the latter case, the FDR is particularly easy to compute because only one Gaussian variable is effective: for step-up procedures,  $\text{FDR}(\text{SU}(\mathbf{t})) = \pi_0 t_m$ ; and for step-down procedures,  $\text{FDR}(\text{SD}(\mathbf{t})) = \sum_{k=1}^m \binom{m}{k} \pi_0^k \pi_1^{m-k} \frac{k}{m} \min\{t_{m-k+1}, \overline{\Phi}(\overline{\Phi}^{-1}(t_1) - \mu)\}$  (the proof is left to the reader). For instance, the above FDR expressions yield  $\text{FDR}(\text{SU}(\mathbf{t})) = t_m$  and  $\text{FDR}(\text{SD}(\mathbf{t})) = t_1$  under the special  $p$ -value configuration where  $\rho = 1$  and  $\pi_0 = 1$ . Thus, as  $\rho \simeq 1$ , the FDR value may considerably change as one consider a step-up or a step-down algorithm.

Going back to Corollary 3.4, let us mention that the latter can be used in order to evaluate the FDR control robustness under Gaussian equicorrelated positive dependence for any procedure (step-up or step-down) that controls the FDR under independence; for instance, the so-called *adaptive* procedures [5] (step-up using  $t_k = \alpha \min\{1, (1 - \alpha)k/(m - k + 1)\}$ ) and [12] (step-down using  $t_k = \alpha k/(m - (1 - \alpha)k + 1)$ ) have been proved to control the FDR at level  $\alpha$  under independence. A simulation study was done in [5] in order to evaluate their robustness in the EMN model with  $\rho > 0$ , that is, in order to see if the FDR is still below  $\alpha$  (or at least close to  $\alpha$ ) in that case. Using our exact approach, we are able to reproduce their analysis without

the errors due to the Monte-Carlo approximation. However, we underline that our approach use *non-random* thresholds  $\mathbf{t}$ ; this is not always the case for adaptive procedures (see e.g. [2]).

3.3. *EMN model with a general correlation and  $m = 2$ .* What happens for  $\rho < 0$ ? The approach presented in the last subsection is not valid anymore in that context and the problem seems considerably more difficult to tackle. We propose in this section to focus on the case where only two hypotheses are tested, which should hopefully give some hints concerning the behavior of the FDR under negative correlations for larger  $m$ . The next result follows from elementary integration and does not require the use of an unconditional model.

PROPOSITION 3.6. *For  $m = 2$  hypotheses, consider the conditional EMN model  $P_{(H,\rho,\mu)}^N$  with parameters  $H = (H_1, H_2) \in \{0, 1\}^2$  (generating  $m_0 \in \{1, 2\}$  true null hypotheses),  $\rho \in [-1, 1]$  and  $\mu > 0$ , and a threshold  $\mathbf{t} = (t_1, t_2)$ . Let  $z_1 = \bar{\Phi}^{-1}(t_2)$  and  $z_2 = \bar{\Phi}^{-1}(t_1)$ . Then  $FDR(SU(\mathbf{t}), P_{(H,\rho,\mu)}^N)$  is given by*

$$(24) \quad \begin{array}{|l|l|} \hline \rho \in (-1, 1), m_0 = 1 & \frac{1}{2} \int_0^{\bar{\Phi}(z_1-\mu)} \bar{\Phi} \left( \frac{z_1 - \rho \bar{\Phi}^{-1}(w)}{\sqrt{1-\rho^2}} \right) dw + \int_{\bar{\Phi}(z_1-\mu)}^1 \bar{\Phi} \left( \frac{z_2 - \rho \bar{\Phi}^{-1}(w)}{\sqrt{1-\rho^2}} \right) dw \\ \hline \rho \in (-1, 1), m_0 = 2 & t_1 + \int_{t_1}^{t_2} \bar{\Phi} \left( \frac{z_1 - \rho \bar{\Phi}^{-1}(w)}{\sqrt{1-\rho^2}} \right) dw + \int_{t_2}^1 \bar{\Phi} \left( \frac{z_2 - \rho \bar{\Phi}^{-1}(w)}{\sqrt{1-\rho^2}} \right) dw \\ \hline \rho = -1, m_0 = 1 & \begin{cases} t_1 & \text{if } 0 < \mu \leq 2z_1 \\ t_1 + \frac{1}{2}t_2 - \frac{1}{2}\bar{\Phi}(\mu - z_1) & \text{if } 2z_1 < \mu < z_1 + z_2 \\ \frac{1}{2}t_2 + \frac{1}{2}\bar{\Phi}(\mu - z_1) & \text{if } \mu \geq z_1 + z_2 \end{cases} \\ \hline \rho = -1, m_0 = 2 & \begin{cases} 2t_1 & \text{if } 1/2 \geq t_2 \\ 2(t_1 + t_2) - 1 & \text{if } 1/2 < t_2, t_1 + t_2 \leq 1 \\ 1 & \text{if } 1/2 < t_2, t_1 + t_2 > 1 \end{cases} \\ \hline \rho = 1, m_0 = 1 & \frac{1}{2}t_2 \\ \rho = 1, m_0 = 2 & t_2 \end{array}$$

and  $FDR(SD(\mathbf{t}), P_{(H,\rho,\mu)}^N)$  is given by

$$(25) \quad \begin{array}{|l|l|} \hline \rho \in (-1, 1), m_0 = 1 & \frac{1}{2} \int_0^{\bar{\Phi}(z_2-\mu)} \bar{\Phi} \left( \frac{z_1 - \rho \bar{\Phi}^{-1}(w)}{\sqrt{1-\rho^2}} \right) dw + \frac{1}{2} \int_{\bar{\Phi}(z_2-\mu)}^{\bar{\Phi}(z_1-\mu)} \bar{\Phi} \left( \frac{z_2 - \rho \bar{\Phi}^{-1}(w)}{\sqrt{1-\rho^2}} \right) dw \\ & + \int_{\bar{\Phi}(z_1-\mu)}^1 \bar{\Phi} \left( \frac{z_2 - \rho \bar{\Phi}^{-1}(w)}{\sqrt{1-\rho^2}} \right) dw \\ \hline \rho \in (-1, 1), m_0 = 2 & t_1 + \int_{t_1}^1 \bar{\Phi} \left( \frac{z_2 - \rho \bar{\Phi}^{-1}(w)}{\sqrt{1-\rho^2}} \right) dw \\ \hline \rho = -1, m_0 = 1 & \begin{cases} t_1 & \text{if } 0 < \mu \leq z_1 + z_2 \\ \frac{1}{2}(t_1 + t_2) - \frac{1}{2}\bar{\Phi}(\mu - z_2) + \frac{1}{2}\bar{\Phi}(\mu - z_1) & \text{if } z_1 + z_2 < \mu < 2z_2 \\ \frac{1}{2}t_2 + \frac{1}{2}\bar{\Phi}(\mu - z_1) & \text{if } \mu \geq 2z_2 \end{cases} \\ \hline \rho = -1, m_0 = 2 & \min(2t_1, 1) \\ \hline \rho = 1, m_0 = 1 & \frac{1}{2} \min(t_2, \bar{\Phi}(z_2 - \mu)) \\ \rho = 1, m_0 = 2 & t_1 \end{array}$$

#### 4. Application to least/most favorable configurations.

4.1. *Least favorable configurations for the FDR.* In order to study the FDR control, one interesting multiple testing issue is to determine which are the value of the parameter  $F_1$  (or  $\mu$ ) for the  $p$ -value distribution that maximizes the value of the FDR. The latter is called a least favorable configuration (LFC) for the FDR.

4.1.1. *Independent model.* Under independence, [3] proved that for any step-up procedure  $SU(\mathbf{t})$  satisfying that  $t_k/k$  is nondecreasing, the LFC is the Dirac-uniform configuration  $F_1 = 1$ . As a matter of fact, they proved more than that: if  $F_1$  and  $F'_1$  are two c.d.f.'s with for all  $t$ ,  $F_1(t) \leq F'_1(t)$ , we then have  $FDR(SU(\mathbf{t}), P_{(H, F_1)}^I) \leq FDR(SU(\mathbf{t}), P_{(H, F'_1)}^I)$  (and thus  $FDR(SU(\mathbf{t}), \bar{P}_{(\pi_0, F_1)}^I) \leq FDR(SU(\mathbf{t}), \bar{P}_{(\pi_0, F'_1)}^I)$  unconditionally; the latter can also be directly derived from (15)). For the LSU, the FDR does not depend on  $F_1$  under independence (it is equal to  $\pi_0\alpha$ ) hence any configuration is a LFC.

An open problem is to determine if the Dirac-uniform configuration is still a LFC for the FDR when using a *step-down* procedure. We propose to address this issue in the case of the LSD procedure, by applying Theorem 3.3. First define the assumption for  $m \geq 2$  and  $\alpha \in (0, 1)$ ,

$$(A(\alpha, m)) \quad k \mapsto \sum_{j=k}^m \frac{k}{j} \binom{m-k}{j-k} \left( \frac{m-\alpha(j+1)}{m-\alpha k} \right)^{m-j} \frac{\alpha}{m-\alpha k} \left( \frac{\alpha(j-k+1)}{m-\alpha k} \right)^{j-k-1} .$$

is a nondecreasing function over  $\{1, \dots, m\}$

Since the above function is discrete, we can easily checked  $A(\alpha, m)$  numerically. In fact, we found that this assumption is true for any  $(\alpha, m) \in \{0.001 + 0.001 * j, 0 \leq j \leq 998\} \times \{2, \dots, 1000\}$ , which may suggest that  $A(\alpha, m)$  holds for any  $(\alpha, m)$ . Unfortunately, we have yet no formal argument to establish the latter. We can now state the following result which is proved in Section 6.5.

COROLLARY 4.1. *For  $m \geq 2$  hypotheses, consider the unconditional independent model  $\bar{P}_{(\pi_0, F_1)}^I$  and the linear step-down procedure LSD. Assume that  $(\alpha, m)$  satisfies  $A(\alpha, m)$ . Then for any  $\pi_0 \in [0, 1]$  and concave c.d.f.  $F_1 \in \mathcal{F}$ , we have*

$$(26) \quad FDR(LSD, \bar{P}_{(\pi_0, F_1)}^I) \leq FDR(LSD, \bar{P}_{(\pi_0, F_1=1)}^I),$$

*meaning that the Dirac-uniform distribution is a least favorable configuration for the FDR of the LSD. Additionally, the upper-bound in (26) is explicitly given in (21).*

4.1.2. *EMN model.* When the  $p$ -values follow the EMN model, [10] conjectured that a LFC for the FDR of the LSU is still the Dirac-uniform distribution (see Section 1 therein). Here, we support this conjecture when  $\rho \geq 0$  but we disprove it when  $\rho < 0$ .

In order to investigate this issue, we reported on Figure 1 the FDR of the LSU procedure against  $\mu$  in the EMN model when  $\rho > 0$  (left) and when  $\rho < 0$  (right), by using Corollary 3.4 and Proposition 3.6. Under positive correlation, although each curve is not necessarily nondecreasing

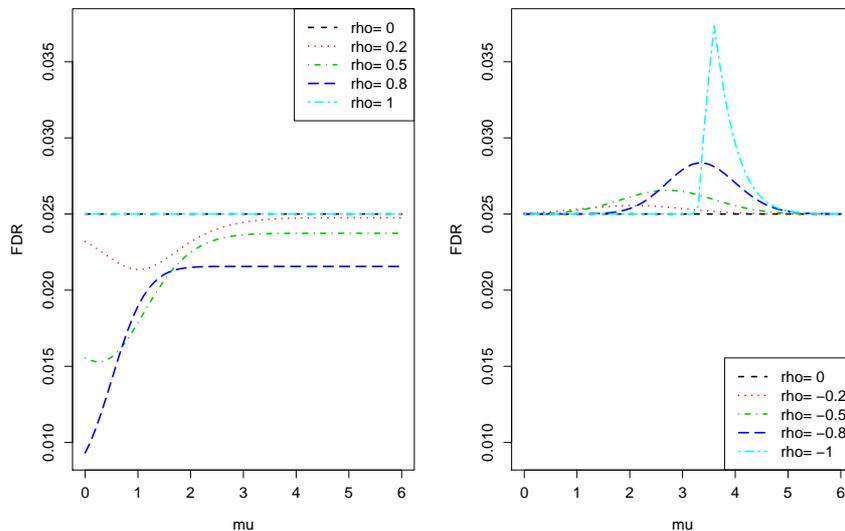


FIG 1.  $FDR(LSU)$  against the mean  $\mu$ . Left:  $\rho \geq 0$  exchangeable EMN model  $m = 100$  and  $\pi_0 = 0.5$ . Right:  $\rho < 0$  conditional model with  $m = 2$  and  $m_0 = 1$ .  $\alpha = 0.05$ .

(e.g. for  $\rho = 0.2$ ), “ $\mu$  very large” seems to be a LFC (we were unfortunately not able to prove the latter formally). Under negative correlation and  $m = 2$ , however,  $\mu$  large is not a LFC anymore. As a matter of fact, in the case where  $\mu = \infty$  and  $m = 2$  (close to the right most point of Figure 1), the two  $p$ -values are independent, with one uniform and the other equal to 0, so that the FDR equals  $\alpha/2 = \alpha m_0/m$  which is not a maximum for the FDR.

Qualitatively, we observed the same behavior concerning the FDR of the LSD procedure.

Under negative correlation, the Dirac-uniform is not a LFC for the FDR and we can therefore legitimately ask what are the LFC’s in that case. Here, we propose to solve this problem when  $m = 2$  in the conditional EMN model. Let  $z_1 = \overline{\Phi}^{-1}(\alpha)$  and  $z_2 = \overline{\Phi}^{-1}(\alpha/2)$  and first consider the LSU procedure. Its FDR is plotted in Figure 2 (top). When  $m_0 = 1$ , we can check that  $(\rho, \mu) = (-1, z_1 + z_2)$  is a LFC because applying (24), the corresponding FDR is  $3\alpha/4$ , which equals the Benjamini-Yekutieli’s (BY) bound  $(1 + 1/2 + \dots + 1/m)\alpha m_0/m$  [3] (valid under any dependency). For  $m_0 = 2$ , we may differentiate the corresponding expression in (24) in order to obtain that, assuming  $\alpha \leq 1/2$ , the FDR (that does not depend on  $\mu$ ) attains its maximum in  $\rho = [-z_1(z_1 - z_2) - \sqrt{(z_1^2 - z_1 z_2)^2 + 2 \log(2)(z_1^2 - z_2^2) + 4 \log(2)}]/(2 \log(2)) \in (-1, 0)$ . Interestingly, [15] states that the BY’s bound can be fulfilled using very specific dependency structures between the  $p$ -values (not necessarily including dependency coming from a EMN model). Here, we remark that the maximum value of the FDR still equals the BY bound for  $m_0 = 1$ , even for  $p$ -values

coming from a EMN model.

Second, we consider the LSD procedure, whose FDR is plotted in Figure 2 (bottom). In the case where  $m_0 = 2$ , by differentiating the corresponding expression in (24) (that does not depend on  $\mu$ ), we are able to state that the FDR attains its maximum at  $\rho = -1$  and that the value of the maximum is  $\alpha$ . In particular, the FDR of the LSD procedure is always smaller than or equal to  $\alpha$  when  $m = 2$ , both in the conditional and unconditional EMN model, even for a negative correlation. An interesting open problem is to know whether this is specific to the case  $m = 2$ .

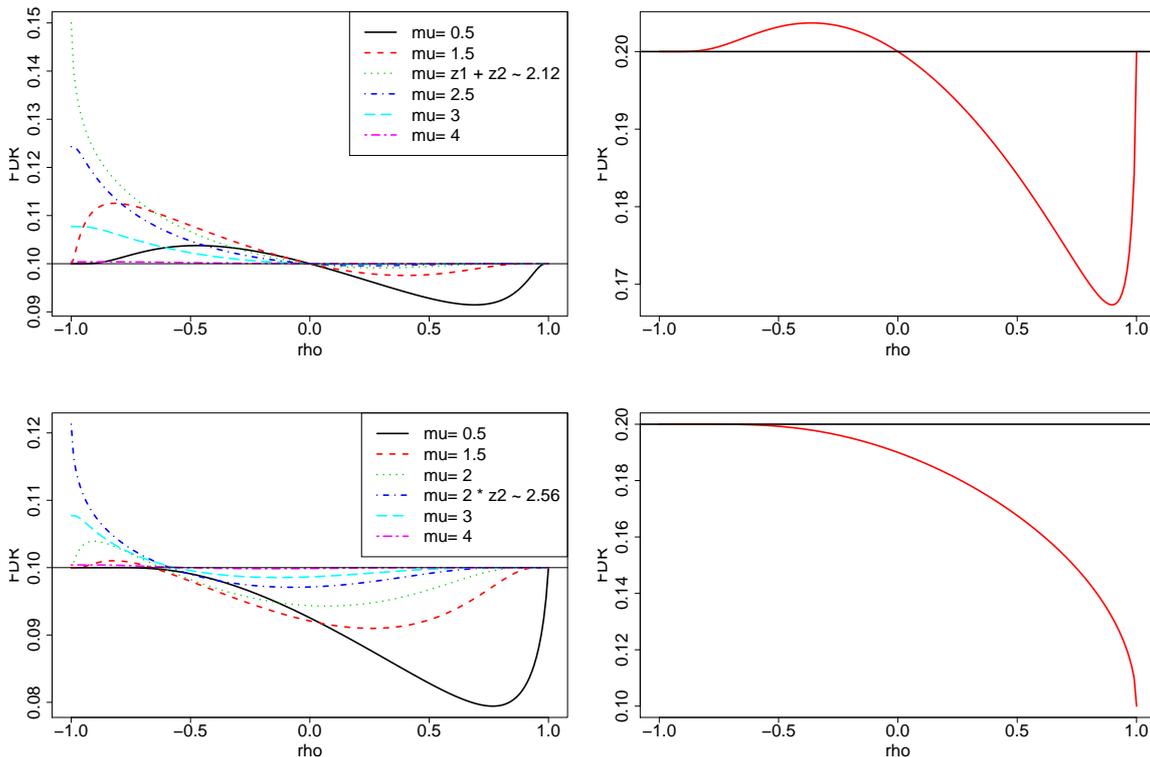


FIG 2. FDR against the covariance  $-1 \leq \rho \leq 1$  in the conditional EMN model.  $m = 2$ ;  $\alpha = 0.2$ . Left:  $m_0 = 1$ , right:  $m_0 = 2$ . Top: LSU, bottom: LSD.

REMARK 4.2. [19] also studied the value of  $\mu$  maximizing the FDR in the case  $m = 2$  in the (conditional) EMN model with possibly negative correlation. The latter work focused on the two-sided testing with  $\rho \in \{-1, 1\}$ ,  $m_0 = 1$  and  $m = 2$ .

4.2. LFC's for the variance of the FDP. We focus here on the unconditional independent model and on the LSU procedure. Using (13) with  $s = 2$ , we easily derive the following expression

for the variance of the FDP:

$$(27) \quad \mathbb{V}[\text{FDP}(\text{LSU})] = \alpha\pi_0 \sum_{k=1}^m \frac{1}{k} \mathcal{D}_{m-1}([G(\alpha(j+1)/m)]_{1 \leq j \leq m-1}, k-1) - (\alpha\pi_0)^2/m.$$

As a consequence, by contrast with the FDR which is constantly equal to  $\pi_0\alpha$  in that case, the variance of the FDP depends on the alternative  $p$ -value c.d.f.  $F_1$ . Moreover, from (27), the smaller  $F_1$ , the larger the variance. Therefore, over the set  $F_1 \in \mathcal{F}$ , the least favorable configuration for the variance (that is, the configuration where the variance is the largest) is given by  $F_1 = 0$  while the most favorable configuration (that is, the configuration where the variance is the smallest) is the Dirac-uniform distribution  $F_1 = 1$ . Over the more “realistic” c.d.f. sets  $\mathcal{F}' = \{F_1 \in \mathcal{F} \mid F_1(x) \geq x\}$  and  $\mathcal{F}_\varepsilon = \{F_1 \in \mathcal{F} \mid F_1(x) \geq \varepsilon\}$ ,  $0 < \varepsilon \leq 1$  the least favorable configurations for the variance are given respectively by  $F_1(x) = x$  and  $F_1(x) = \varepsilon$ . For these extreme configurations, the expression of the variance (27) can be simplified by using the next formula (proved in Section 6.6): for any threshold of the form  $t_k = \beta + k\gamma$ ,  $1 \leq k \leq m$ , with  $\beta, \gamma \geq 0$ ,

$$(28) \quad \sum_{k=0}^m \frac{1}{k+1} \mathcal{D}_m(\mathbf{t}, k) = \frac{1}{\gamma - \beta} \left[ \frac{(1 + \gamma - \beta)^{m+1} - 1}{m+1} - \gamma[(1 + \gamma - \beta)^m - 1] \right]$$

for  $\gamma \neq \beta$  and  $\sum_{k=0}^m \frac{1}{k+1} \mathcal{D}_m(\mathbf{t}, k) = 1 - m\gamma$  otherwise. This leads to the following results.

**COROLLARY 4.3.** *Consider the linear step-up procedure LSU in the unconditional independent model with parameters  $\pi_0, F_1$ . Then for any  $m \geq 2$ ,  $\alpha \in (0, 1)$  and  $\pi_0 \in [0, 1]$ , under the generating distribution  $(H, \mathbf{p}) \sim \bar{P}_{(\pi_0, F_1)}^I$ , the following lower-bounds hold:*

$$(29) \quad \begin{aligned} \min_{F_1 \in \mathcal{F}} \{\mathbb{V}[\text{FDP}(\text{LSU}, H)]\} &= \min_{F_1 \in \mathcal{F}'} \{\mathbb{V}[\text{FDP}(\text{LSU}, H)]\} = \min_{F_1 \in \mathcal{F}_\varepsilon} \mathbb{V}[\text{FDP}(\text{LSU}, H)] \\ &= \frac{\alpha\pi_0}{m} \frac{1 - \pi_0^m}{1 - \pi_0} - \frac{(\alpha\pi_0)^2}{m} \left( \frac{1 - \pi_0^{m-1}}{1 - \pi_0} + 1 \right). \end{aligned}$$

Moreover, the following upper-bounds hold:

$$\begin{aligned} \max_{F_1 \in \mathcal{F}} \{\mathbb{V}[\text{FDP}(\text{LSU}, H)]\} &= \alpha\pi_0(1 - \alpha\pi_0) \\ \max_{F_1 \in \mathcal{F}'} \{\mathbb{V}[\text{FDP}(\text{LSU}, H)]\} &= \alpha\pi_0(1 - \alpha) + (1 - \pi_0) \frac{\pi_0\alpha^2}{m} \\ \max_{F_1 \in \mathcal{F}_\varepsilon} \{\mathbb{V}[\text{FDP}(\text{LSU}, H)]\} &= \frac{\alpha\pi_0}{m} \frac{1 - (1 - (1 - \pi_0)\varepsilon)^m}{(1 - \pi_0)\varepsilon} - \frac{(\alpha\pi_0)^2}{m} \left( \frac{1 - (1 - (1 - \pi_0)\varepsilon)^{m-1}}{(1 - \pi_0)\varepsilon} + 1 \right). \end{aligned}$$

The proof is made in Section 6.6.

Using Corollary 4.3, we are able to investigate the following asymptotic multiple testing issue: does the FDP converge to the FDR as  $m$  grows? Establishing the latter is crucial because when

one establishes  $\text{FDR} \leq \gamma$ , one implicitly want that for the observed realization  $\omega$  we still have  $\text{FDP}(\omega) \leq \gamma$ . Here, the variance measure the  $L^2$  distance between these two quantities, and since  $\text{FDP} \in [0, 1]$  the latter distance tends to zero if and only if the FDP converges to the FDR in probability. First, if  $\pi_0 \in (0, 1)$  is fixed with  $m$ , the convergence holds over the set of c.d.f.'s  $\mathcal{F}_\varepsilon$ , with a distance  $\mathbb{E}(\text{FDP} - \text{FDR})^2$  converging to zero at rate  $1/m$ . This corroborate previous asymptotic studies in the so-called ‘‘sub-critical’’ case (see e.g. [6, 17]). By contrast, when considering the classes  $\mathcal{F}$  and  $\mathcal{F}'$  the convergence does not hold in the least favorable configurations  $F_1(x) = 0$  and  $F_1(x) = x$ , respectively. The latter is quite intuitive because the denominator in the FDP does not converges to infinity anymore in that cases (see e.g. [11]), so these configurations can probably be considered as marginal. Second, our non-asymptotic approach allows to make  $\pi_0$  depends on  $m$  and to explore the case where  $1 - \pi_0 = 1 - \pi_{0,m} \sim m^{-\beta}$  with  $0 < \beta \leq 1$ , which corresponds to a classical ‘‘sparse’’ setting (see e.g. [7]). Expression (29) implies that, in this sparse case, the variance is always larger than a quantity of order  $1/m^{1-\beta}$ . In particular, when  $1 - \pi_{0,m} \sim 1/m$ , for any  $F_1$ , the FDP does not converge to the FDR, and when  $1 - \pi_{0,m} \sim m^{-\beta}$  with  $0 < \beta < 1$ , for all  $F_1$ , the convergence of the FDP towards the FDR is of order slower than  $1/m^{(1-\beta)/2}$  (in  $L^2$  norm). As illustration, for  $m = 10000$ ,  $1 - \pi_0 = 1/100$  and  $\alpha = 0.05$ , (29) gives that  $(\mathbb{E}(\text{FDP} - \text{FDR})^2)^{1/2} \simeq 0.0217$ , so the FDP has a distribution quite spread around the  $\text{FDR} = \pi_0\alpha \simeq 0.05$ . As a conclusion, considering a sparse signal slow down the convergence of the FDP to the FDR, so any FDR control should be interpreted with cautious in such a situation.

**5. Discussion.** Several extensions of this work can be proposed. First, in the independent model, we may replace in our formulas  $F_0(x) = x$  by any c.d.f. For instance, this can be useful to consider  $p$ -values coming from discrete test statistics, for which the case  $F_0(x) < x$  is possible. Second, as described in Section 2 of [10], the trick that we used in Section 3.2 to directly extend our formulas from the independent case to the equi-correlated normal case works for any suitable test statistic of the form  $T_i = g(X_i, U)$  with independent variables  $(X_i)_i$  and a variable  $U$  independent from the  $X_i$ 's. Therefore, our methodology also applied in such a model. Third, our approach can be used in order to compute exactly the positive false discovery rate  $\text{pFDR}(R) = \text{FDR}(R)/\mathbb{P}[|R| > 0]$  (as defined in [26]) and the false non-discovery rate FNR, defined as the expected proportion of false hypotheses among the non-rejected hypotheses. For instance, for the FNR, in the exchangeable independent model, we may prove that

$$\text{FNR}(\text{SD}(\mathbf{t}), \bar{P}_{\pi_0, F}^I) = m\pi_1 \sum_{k=0}^{m-1} \frac{1 - F_1(t_{k+1})}{m - k} \tilde{\mathcal{D}}_{m-1}((G(t_j))_{1 \leq j \leq m-1}, k).$$

For the FNR of a step-up procedure  $\text{SU}(\mathbf{t})$  rejecting  $\hat{k}$  hypotheses, the exact expression involves the joint distribution between  $\hat{k}_{(1)}$  and  $\hat{k}'_{(1)}$  (with definitions analogue to Lemma A.2). As a matter of fact, it is interesting to note the following duality in our exact expressions: the FDR of a step-up (resp. step-down) procedure has a form similar to the FNR of a step down (resp. step-

up) procedure. This is related to the following property: point-wise, the hypotheses rejected by  $\text{SD}(\mathbf{t}, \mathbf{p})$  are exactly the hypotheses non-rejected by  $\text{SU}(\bar{\mathbf{t}}, \bar{\mathbf{p}})$  with  $\bar{p}_i = 1 - p_i$  and  $\bar{t}_r = 1 - t_{m-r+1}$ .

Furthermore, since our formulas depend on the true parameters of the model, which are in general unknown in a statistical approach, one may formulate the concern of estimating these quantities in our formulas. We did not investigate in detail this issue as it would exceed the scope of this paper. Here, we simply notice that plugging convergent estimators of the parameters in our formulas will lead to convergent estimators for the corresponding quantity (e.g. FDR, Power), because our formulas are continuous in all the model parameters.

Finally, let us underline that the bounds for the variance of the FDP obtained in Corollary 4.3 show that the FDR can be sometimes far to the real observed FDP value, even for a standard setting assuming independent  $p$ -values and considering the linear step-up procedure. This reinforces the interest in FDP controls of the form  $\mathbb{P}[\text{FDP} \leq \alpha] \geq 1 - \gamma$ , as suggested in [13, 16].

## 6. Proofs.

6.1. *Proof of Theorem 3.1.* The proof uses an argument similar to [9] (including the correction by the coefficients  $\left\{ \binom{s}{\ell} \right\}$ ). By definition of a step-up procedure, we have  $\text{SU}(\mathbf{t}) = \{i \mid p_i \leq t_{\hat{k}}\}$  with  $\hat{k} = |\text{SU}(\mathbf{t})|$  (expression related to the ‘‘self-consistency’’ condition introduced in [4]). Thus, when  $(H, \mathbf{p}) \sim \bar{P}_{(\pi_0, F_1)}^I$ , the  $s$ -th moment of the FDP can be written as

$$\mathbb{E}[\text{FDP}(\text{SU}(\mathbf{t}), H)^s] = \mathbb{E} \left[ (1 \vee \hat{k}(\mathbf{p}))^{-s} \left( \sum_{i=1}^m \mathbf{1}\{H_i = 0, p_i \leq t_{\hat{k}(\mathbf{p})}\} \right)^s \right]$$

Letting  $X_i = \mathbf{1}\{p_i \leq t_{\hat{k}(\mathbf{p})}\}$  and  $Y_i = \mathbf{1}\{H_i = 0\}$ , the variables couples  $(X_1, Y_1), \dots, (X_m, Y_m)$  are exchangeable because of the exchangeability of  $(H_i, p_i)_{1 \leq i \leq m}$  and because  $\hat{k}(\cdot)$  has exchangeable coordinates. Moreover, the function  $f$  defined as  $f(X_1, \dots, X_m) = ((X_1 + \dots + X_m) \vee 1)^{-1}$  has also exchangeable coordinates. As a consequence, we may apply Lemma A.3, to obtain

$$\mathbb{E}[\text{FDP}(\text{SU}(\mathbf{t}), H)^s] = \sum_{\ell=1}^s \frac{m!}{(m-\ell)!} \left\{ \binom{s}{\ell} \right\} \mathbb{E} \left[ (1 \vee \hat{k})^{-s} \prod_{1 \leq i \leq \ell} \mathbf{1}\{H_i = 0, p_i \leq t_{\hat{k}}\} \right].$$

Next, using Lemma A.1 (and the notation therein), combined with the independence between the  $p$ -values, we obtain

$$\begin{aligned} \mathbb{E}[\text{FDP}(\text{SU}(\mathbf{t}), H)^s] &= \sum_{\ell=1}^s \frac{m!}{(m-\ell)!} \binom{s}{\ell} \mathbb{E} \left[ (\hat{k}'_{(\ell)} + \ell)^{-s} \prod_{1 \leq i \leq \ell} \mathbf{1}\{H_i = 0, p_i \leq t_{\hat{k}'_{(\ell)} + \ell}\} \right] \\ &= \sum_{\ell=1}^s \frac{m!}{(m-\ell)!} \binom{s}{\ell} \mathbb{E} \left[ (\hat{k}'_{(\ell)} + \ell)^{-s} \mathbb{E} \left[ \prod_{1 \leq i \leq \ell} \mathbf{1}\{H_i = 0, p_i \leq t_{\hat{k}'_{(\ell)} + \ell}\} \middle| \hat{k}'_{(\ell)} \right] \right] \\ &= \sum_{\ell=1}^s \frac{m!}{(m-\ell)!} \binom{s}{\ell} \pi_0^\ell \mathbb{E} \left[ \frac{(F_0(t_{\hat{k}'_{(\ell)} + \ell}))^\ell}{(\hat{k}'_{(\ell)} + \ell)^s} \right]. \end{aligned}$$

Combining the latter expression with (12) leads to (13).

The power computation uses analogue technics:  $\text{Pow}(\text{SU}(\mathbf{t})) = m^{-1} \sum_{i=1}^m \mathbb{P}[p_i \leq t_{\hat{k}} \mid H_i = 1] = \mathbb{P}[p_1 \leq t_{\hat{k}} \mid H_1 = 1] = \mathbb{P}[p_1 \leq t_{\hat{k}_{(1)}+1} \mid H_1 = 1] = \mathbb{E}[F_1[t_{\hat{k}_{(1)}+1}]]$ .

*6.2. Proof of Theorem 3.3.* To show (19), we give a proof analogue to the one of Theorem 3.1, except that we use Lemma A.2 instead of Lemma A.1. Define  $\tilde{k} = |\text{SD}(\mathbf{t})|$  and  $\tilde{k}_{(1)}, \tilde{k}'_{(1)}$  as in Lemma A.2. We get by exchangeability of  $(H_i, p_i)_i$  and independence of the  $p$ -values,

$$\begin{aligned} \text{FDR}(R) &= \sum_{i=1}^m \mathbb{E} \left[ \frac{\mathbf{1}\{p_i \leq t_{\tilde{k}}\} \mathbf{1}\{H_i = 0\}}{\tilde{k} \vee 1} \right] = m \mathbb{E} \left[ \frac{\mathbf{1}\{p_1 \leq t_{\tilde{k}}\} \mathbf{1}\{H_1 = 0\}}{\tilde{k} \vee 1} \right] \\ &= m \mathbb{E} \left[ \frac{\mathbf{1}\{p_1 \leq t_{\tilde{k}_{(1)}+1}\} \mathbf{1}\{H_1 = 0\}}{\tilde{k}'_{(1)} + 1} \right] = \pi_0 m \mathbb{E} \left[ \frac{F_0(t_{\tilde{k}_{(1)}+1})}{\tilde{k}'_{(1)} + 1} \right]. \end{aligned}$$

The result follows. The proof for the power is similar, except that it only uses the distribution of  $\tilde{k}_{(1)}$ .

*6.3. Proof of Theorem 3.6.* We focus on step-up procedures, in the case  $\rho \in (-1, 1)$  and  $m_0 = 1$  (the remaining cases are left to the reader). Without loss of generality, we may assume that the first coordinate correspond to the true null, that is,  $H = (0, 1)$ . In this context, the FDP takes one of the three values:  $0, \frac{1}{2}, 1$ , according to the location of the tests statistics  $Y_i = X_i + \mu H_i$  with respect to the critical values  $z_1$  and  $z_2$ . From the definition of a step-up procedure, we may define the two regions for  $i \in \{1, 2\}$ ,  $\mathcal{D}_i = \{(y_1, y_2) \in \mathbb{R}^2 \mid \text{FDP}(y_1, y_2) = i/2\}$ , where  $\text{FDP}(y_1, y_2)$  denotes the FDP of  $\text{SU}(\mathbf{t})$  taken in the  $p$ -values  $\mathbf{p} = (\overline{\Phi}(y_1), \overline{\Phi}(y_2))$ . The regions  $\mathcal{D}_i$  are represented on Figure 3(a). Next, since  $(Y_1, Y_2)$  follows the EMN model, we may write for

$i \in \{1, 2\}$ ,

$$\begin{aligned} \mathbb{P}(\text{FDP}(Y_1, Y_2) = i/2) &= \frac{1}{2\pi\sqrt{1-\rho^2}} \int \int_{\mathcal{D}_i} e^{-\frac{1}{2(1-\rho^2)}(y_1 - \rho(y_2 - \mu))^2 - \frac{1}{2}(y_2 - \mu)^2} dy_1 dy_2 \\ &= \frac{1}{2\pi} \int \int_{\tilde{\mathcal{D}}_i} e^{-\frac{1}{2}(u^2 + v^2)} dudv \end{aligned}$$

by using the substitution  $u = \frac{1}{\sqrt{1-\rho^2}}(y_1 - \rho(y_2 - \mu))$  and  $v = y_2 - \mu$ , and where the resulting integration domain  $\tilde{\mathcal{D}}_i$  is represented on Figure 3(b). Therefore, we finally obtain

$$\mathbb{P}(\text{FDP} = 1/2) = \int_{z_1 - \mu}^{\infty} \frac{e^{-\frac{1}{2}v^2}}{\sqrt{2\pi}} \bar{\Phi}\left(\frac{z_1 - \rho v}{\sqrt{1-\rho^2}}\right) dv, \quad \mathbb{P}(\text{FDP} = 1) = \int_{-\infty}^{z_1 - \mu} \frac{e^{-\frac{1}{2}v^2}}{\sqrt{2\pi}} \bar{\Phi}\left(\frac{z_2 - \rho v}{\sqrt{1-\rho^2}}\right) dv,$$

and the final expression results by using the substitution  $w = \bar{\Phi}(v)$ .

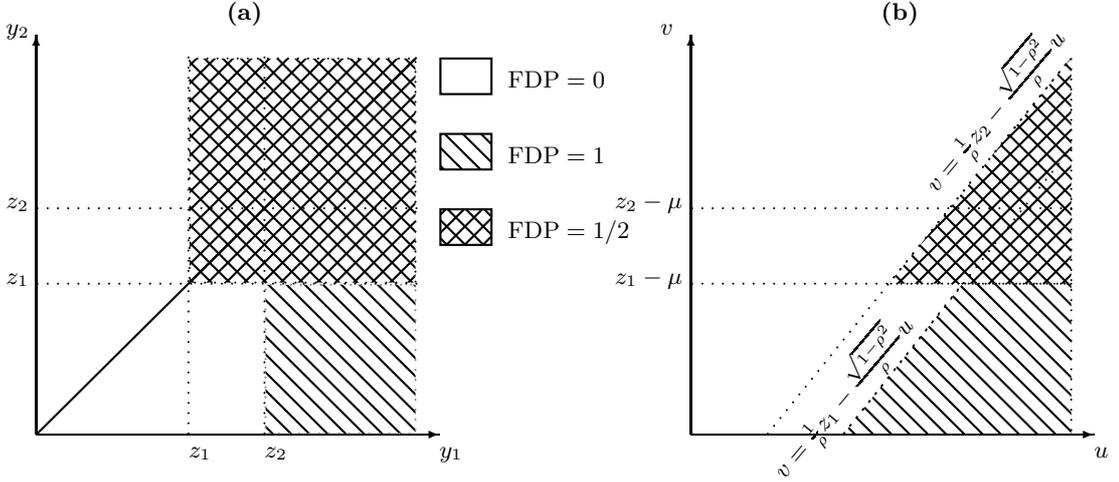


FIG 3. Left:  $\mathcal{D}_i$ . Right:  $\tilde{\mathcal{D}}_i$ . For  $i = 1$  (double cross-hatch) and  $i = 2$  (simple cross-hatch), see text.

6.4. *Proof of (18)*. First remark that we may assumed that  $G(x) = x$  up to replace  $\mathbf{t} = (t_j)_j$  by  $(G(t_j))_j$ . Next, assume  $k < k' < m$  and denote  $L(r) = \sum_{i=1}^m \mathbf{1}\{p_i \leq t_r\}$ . By definition of a

step-down procedure, the probability  $\mathbb{P}[|\text{SD}(\mathbf{t})| = k, |\text{SD}(\mathbf{t}')| = k']$  is equal to

$$\begin{aligned}
& \mathbb{P}[\forall j \leq k, L(j) \geq j, L(k+1) = k, \forall j, k+1 \leq j \leq k', L(j+1) \geq j, L(k'+2) = k'] \\
&= \binom{m}{k} \binom{m-k}{k-k'} \mathbb{P}[\forall j \leq k, L(j) \geq j, \forall j, k+1 \leq j \leq k', L(j+1) \geq j, \\
& \quad p_1, \dots, p_k \leq t_{k+1} < p_{k+1}, \dots, p_{k'} \leq t_{k'+2} < p_{k'+1}, \dots, p_m] \\
&= \binom{m}{k} \binom{m-k}{k-k'} \mathbb{P}\left[\forall j \leq k, \sum_{i=1}^k \mathbf{1}\{p_i \leq t_j\} \geq j, \forall j, k+1 \leq j \leq k', k + \sum_{i=k+1}^{k'} \mathbf{1}\{p_i \leq t_{j+1}\} \geq j, \right. \\
& \quad \left. p_1, \dots, p_k \leq t_{k+1} < p_{k+1}, \dots, p_{k'} \leq t_{k'+2} < p_{k'+1}, \dots, p_m\right] \\
&= \binom{m}{k} \binom{m-k}{k-k'} \mathbb{P}[p_{(1)} \leq t_1, \dots, p_{(k)} \leq t_k] (1 - t_{k'+2})^{m-k} \\
& \quad \mathbb{P}[t_{k+1} < p_{(1)} \leq t_{k+2}, \dots, p_{(k'-k)} \leq t_{k'+1}],
\end{aligned}$$

where we used again that the  $p$ -values are exchangeable. Simple computations give that

$$\mathbb{P}[t_{k+1} < p_{(1)} \leq t_{k+2}, \dots, p_{(k'-k)} \leq t_{k'+1}] = \mathbb{P}[p_{(1)} \leq t_{k+2} - t_{k+1}, \dots, p_{(k'-k)} \leq t_{k'+1} - t_{k+1}].$$

This leads to (18). The cases  $k < k' = m$  and  $k = k'$  are similar.

6.5. *Proof of Corollary 4.1*. For any  $t \in [0, 1]$ , let  $G(t) = \pi_0 t + \pi_1 F(t)$  and  $G_1(t) = \pi_0 t + \pi_1$ . First, since  $F_1$  is concave, we have for  $t < t' \leq 1$ , that  $(F(t') - F(t))/(t' - t) \geq (1 - F(t))/(1 - t)$  and thus

$$(30) \quad (G(t') - G(t))/(1 - G(t)) \geq (G_1(t') - G_1(t))/(1 - G_1(t)).$$

We denote  $\underline{\mathbf{t}} = (\alpha j/m)_{1 \leq j \leq m-1}$ ,  $\bar{\mathbf{t}} = (\alpha j/m)_{2 \leq j \leq m}$ ,  $G \circ \underline{\mathbf{t}} = (G(\alpha j/m))_{1 \leq j \leq m-1}$  and  $G \circ \bar{\mathbf{t}} = (G(\alpha j/m))_{2 \leq j \leq m}$ . From (19) and (10), we obtain that

$$\begin{aligned}
\text{FDR}(LSD, \bar{P}_{(\pi_0, F)}^I) &= \pi_0 m \sum_{k=1}^m \tilde{\mathcal{D}}_{m-1}(G \circ \underline{\mathbf{t}}, k-1) \sum_{i=0}^{m-k} \frac{k}{k+i} \tilde{\mathcal{D}}_{m-k} \left( \left( \frac{G(t_{k+j}) - G(t_k)}{1 - G(t_k)} \right)_{1 \leq j \leq m-k}, i \right) \\
&\leq \pi_0 m \sum_{k=1}^m \tilde{\mathcal{D}}_{m-1}(G \circ \underline{\mathbf{t}}, k-1) \sum_{i=0}^{m-k} \frac{k}{k+i} \tilde{\mathcal{D}}_{m-k} \left( \left( \frac{G_1(t_{k+j}) - G_1(t_k)}{1 - G_1(t_k)} \right)_{1 \leq j \leq m-k}, i \right) \\
&= \pi_0 m \sum_{k=1}^m \tilde{\mathcal{D}}_{m-1}(G \circ \underline{\mathbf{t}}, k-1) \sum_{i=0}^{m-k} \frac{k}{k+i} \tilde{\mathcal{D}}_{m-k} \left( \left( \frac{\alpha j/m}{1 - \alpha k/m} \right)_{1 \leq j \leq m-k}, i \right),
\end{aligned}$$

where the inequality comes from (30) and because for a fixed  $k$ , the sum over  $i$  can be seen as the expectation of  $k/(k+I)$  where  $I$  is the rejection number of a step-down procedure

(point-wise nondecreasing in the threshold). Next, considering this time the sum over  $k$  as an expectation, since  $G \leq G_1$  and since a step-down procedure is point-wise nondecreasing in the threshold, the proof is finished if we prove that  $k \mapsto \sum_{i=0}^{m-k} \frac{k}{k+i} \tilde{\mathcal{D}}_{m-k} \left( \left( \frac{\alpha j/m}{1-\alpha k/m} \right)_{1 \leq j \leq m-k}, i \right)$  is nondecreasing in  $k$ . By using (9) and (7), this is assumption  $A(\alpha, m)$ .

6.6. *Proofs for Section 4.2.* Let us first prove (28): denote for any  $0 \leq i \leq m-1$ ,  $u_i = \mathbb{E}[ (|\text{SU}((t_j)_{i+1 \leq j \leq m})| + i + 1)^{-1} ]$  and  $u_m = 1/(m+1)$ ,  $u_{m+1} = 0$ , so that  $u_0$  equals the quantity in (28). We may prove the following recursion relation: for any  $0 \leq i \leq m$ ,

$$(31) \quad (i+1)u_i = (1 - (m-i)\gamma) - (m-i)(\beta - \gamma)u_{i+1}.$$

Expression (31) is proved as follows: for  $i < m$ ,

$$\begin{aligned} (i+1)u_i &= 1 - \mathbb{E} \left[ \frac{|\text{SU}((t_j)_{i+1 \leq j \leq m})|}{|\text{SU}((t_j)_{i+1 \leq j \leq m})| + i + 1} \right] \\ &= 1 - \sum_{k=1}^{m-i} \frac{k}{k+i+1} \binom{m-i}{k} (t_{k+i})^k \Psi_{m-i-k}(1-t_m, \dots, 1-t_{k+i+1}) \\ &= 1 - (m-i) \sum_{k=1}^{m-i} \frac{t_{k+i}}{k+i+1} \binom{m-i-1}{k-1} (t_{k+i})^{k-1} \Psi_{m-i-k}(1-t_m, \dots, 1-t_{k+i+1}) \\ &= 1 - (m-i) \sum_{k=0}^{m-(i+1)} \left( \gamma + \frac{\beta - \gamma}{k + (i+1) + 1} \right) \binom{m-(i+1)}{k} (t_{k+(i+1)})^k \\ &\quad \times \Psi_{m-(i+1)-k}(1-t_m, \dots, 1-t_{k+(i+1)+1}). \end{aligned}$$

Next, we obtain that the solution of the recursion (31) is given by

$$u_i = \sum_{j=0}^{m-i} \frac{1 - (m-(i+j))\gamma}{m-(i+j)} \frac{(m-i) \times \dots \times (m-(i+j))}{(i+1) \times \dots \times (i+j+1)} (\gamma - \beta)^j,$$

which leads to  $u_0 = \sum_{j=0}^m \frac{1-(m-j)\gamma}{j+1} \binom{m}{j} (\gamma - \beta)^j$  and (28) results.

To prove Corollary 4.3, we combine (27) and (28), the latter using  $m-1$  hypotheses and special values for  $\beta$  and  $\gamma$ :  $\beta = \gamma = \pi_0 \alpha/m$  for  $F_1(x) = 0$ ;  $\beta = \gamma = \alpha/m$  for  $F_1(x) = x$ ;  $\beta = \pi_0 \alpha/m + (1 - \pi_0)\varepsilon$  and  $\gamma = \pi_0 \alpha/m$  for  $F_1(x) = \varepsilon$ .

**Acknowledgements.** We would like to warmly acknowledge Gilles Blanchard, Sylvain Delattre and Zhan Shi for helpful discussions. This work was supported by the French Agence Nationale de la Recherche (ANR grant references: ANR-09-JCJC-0027-01, ANR-PARCIMONIE, ANR-09-JCJC-0101-01).

## References.

- [1] Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B*, 57(1):289–300.
- [2] Benjamini, Y., Krieger, A. M., and Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, 93(3):491–507.
- [3] Benjamini, Y. and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Ann. Statist.*, 29(4):1165–1188.
- [4] Blanchard, G. and Roquain, E. (2008). Two simple sufficient conditions for FDR control. *Electron. J. Stat.*, 2:963–992.
- [5] Blanchard, G. and Roquain, E. (2009). Adaptive FDR control under independence and dependence. *J. Mach. Learn. Res.*, 10:2837–2871.
- [6] Chi, Z. (2007). On the performance of FDR control: constraints and a partial solution. *Ann. Statist.*, 35(4):1409–1431.
- [7] Donoho, D. and Jin, J. (2004). Higher criticism for detecting sparse heterogeneous mixtures. *Ann. Statist.*, 32(3):962–994.
- [8] Efron, B., Tibshirani, R., Storey, J. D., and Tusher, V. (2001). Empirical Bayes analysis of a microarray experiment. *J. Amer. Statist. Assoc.*, 96(456):1151–1160.
- [9] Ferreira, J. A. and Zwinderman, A. H. (2006). On the Benjamini-Hochberg method. *Ann. Statist.*, 34(4):1827–1849.
- [10] Finner, H., Dickhaus, T., and Roters, M. (2007). Dependency and false discovery rate: asymptotics. *Ann. Statist.*, 35(4):1432–1455.
- [11] Finner, H. and Roters, M. (2002). Multiple hypotheses testing and expected number of type I errors. *Ann. Statist.*, 30(1):220–238.
- [12] Gavrilov, Y., Benjamini, Y., and Sarkar, S. K. (2009). An adaptive step-down procedure with proven fdr control under independence. *Ann. Statist.*, 37(2):619–629.
- [13] Genovese, C. and Wasserman, L. (2004). A stochastic process approach to false discovery control. *Ann. Statist.*, 32(3):1035–1061.
- [14] Glueck, D., Mandel, J., Karimpour-Fard, A., Hunter, L., and Muller, K. (2008). Exact calculations of average power for the benjamini-hochberg procedure. *International Journal of Biostatistics*, 4(1):1103–1103.
- [15] Guo, W. and Rao, M. B. (2008). On control of the false discovery rate under no assumption of dependency. *Journal of Statistical Planning and Inference*, 138(10):3176–3188.
- [16] Lehmann, E. L. and Romano, J. P. (2005). Generalizations of the familywise error rate. *Ann. Statist.*, 33:1138–1154.
- [17] Neuvial, P. (2008). Asymptotic properties of false discovery rate controlling procedures under independence. *Electron. J. Stat.*, 2:1065–1110.
- [18] Owen, D. B. and Steck, G. P. (1962). Moments of order statistics from the equicorrelated multivariate normal distribution. *Ann. Math. Statist.*, 33:1286–1291.
- [19] Reiner-Benaim, A. (2007). FDR control by the BH procedure for two-sided correlated tests with implications to gene expression data analysis. *Biom. J.*, 49(1):107–126.
- [20] Roquain, E. and van de Wiel, M. (2009). Optimal weighting for false discovery rate control. *Electron. J. Stat.*, 3:678–711.
- [21] Sarkar, S. K. (2002). Some results on false discovery rate in stepwise multiple testing procedures. *Ann. Statist.*, 30(1):239–257.
- [22] Sarkar, S. K. (2008). On methods controlling the false discovery rate. *Sankhya, Ser. A*, 70:135–168.
- [23] Shorack, G. R. and Wellner, J. A. (1986). *Empirical processes with applications to statistics*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York.
- [24] Simes, R. J. (1986). An improved Bonferroni procedure for multiple tests of significance. *Biometrika*, 73(3):751–754.

- [25] Storey, J. D. (2002). A direct approach to false discovery rates. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 64(3):479–498.
- [26] Storey, J. D. (2003). The positive false discovery rate: a Bayesian interpretation and the  $q$ -value. *Ann. Statist.*, 31(6):2013–2035.
- [27] Stuart, A. (1958). Equally correlated variates and the multinormal integral. *J. Roy. Statist. Soc. Ser. B*, 20:373–378.

## Appendix

### APPENDIX A: USEFUL LEMMAS

The following lemma is related to the proof of Theorem 2.1 in [9] and to Lemma 8.1 (i) of [20].

LEMMA A.1. *Consider a step-up procedure  $SU(\mathbf{t})$  using a given threshold  $\mathbf{t}$  testing  $m$  null hypotheses and rejecting  $\hat{k} = |SU(\mathbf{t})|$  hypotheses. For a given  $1 \leq \ell \leq m$ , denote by  $\hat{k}'_{(\ell)}$  the number of rejections of the step-up procedure of threshold  $(t_{j+\ell})_{1 \leq j \leq m-\ell}$  over the  $m - \ell$  last hypotheses. Then we have point-wise*

$$\forall 1 \leq i \leq \ell, p_i \leq t_{\hat{k}} \iff \forall 1 \leq i \leq \ell, p_i \leq t_{\hat{k}'_{(\ell)} + \ell} \iff \hat{k} = \hat{k}'_{(\ell)} + \ell.$$

PROOF. First note that since  $\hat{k}(\cdot)$  is nondecreasing in each coordinates, we always have  $\hat{k} \leq \hat{k}'_{(\ell)} + \ell$ . Second, since  $p_{(k)} \leq t_k$  is equivalent to  $|\{1 \leq j \leq m \mid p_j \leq t_k\}| \geq k$  the rejection number of a step-up procedure can be defined as  $\hat{k} = \max\{k \in \{0, 1, \dots, m\} \mid |\{1 \leq j \leq m \mid p_j \leq t_k\}| \geq k\}$ . Hence,  $\forall 1 \leq i \leq \ell, p_i \leq t_{\hat{k}'_{(\ell)} + \ell}$  is equivalent to  $|\{1 \leq j \leq m \mid p_j \leq t_{\hat{k}'_{(\ell)} + \ell}\}| \geq \ell + |\{ \ell + 1 \leq j \leq m \mid p_j \leq t_{\hat{k}'_{(\ell)} + \ell}\}|$  which is equivalent to  $|\{1 \leq j \leq m \mid p_j \leq t_{\hat{k}'_{(\ell)} + \ell}\}| \geq \ell + \hat{k}'_{(\ell)}$  by definition of  $\hat{k}'_{(\ell)}$ . As a consequence, since  $\hat{k}$  is a maximum and since  $\hat{k} \leq \hat{k}'_{(\ell)} + \ell$ , the latter is equivalent to  $\hat{k}'_{(\ell)} + \ell = \hat{k}$ . This establishes the second equivalence. The first equivalence easily comes from the second equivalence and using that  $t_{\hat{k}} \leq t_{\hat{k}'_{(\ell)} + \ell}$  because  $(t_k)_k$  is a nondecreasing sequence.  $\square$

For step-down procedures, we should use the next lemma.

LEMMA A.2. *Consider a step-down procedure  $SD(\mathbf{t})$  using a given threshold  $\mathbf{t}$  testing  $m$  null hypotheses and rejecting  $\tilde{k} = |SD(\mathbf{t})|$  hypotheses. Denote by  $\tilde{k}_{(1)}$  (resp.  $\tilde{k}'_{(1)}$ ) the number of rejections of the step-up procedure of threshold  $(t_j)_{1 \leq j \leq m-1}$  (resp.  $(t_{j+1})_{1 \leq j \leq m-1}$ ) over the  $m - 1$  last hypotheses. Then we have point-wise*

$$p_1 \leq t_{\tilde{k}} \iff p_1 \leq t_{\tilde{k}_{(1)} + 1} \iff \tilde{k} = \tilde{k}'_{(1)} + 1.$$

In the above lemma, the assertion  $p_1 \leq t_{\tilde{k}_{(1)} + 1} \implies p_1 \leq t_{\tilde{k}}$  is not true in general.

PROOF. Similarly to the step-up case, the rejection number of a step-down procedure can be defined as  $\tilde{k} = \max\{k \in \{0, 1, \dots, m\} \mid \forall k' \leq k, |\{1 \leq j \leq m \mid p_j \leq t_{k'}\}| \geq k'\}$ . Also remark that we always have  $\tilde{k}'_{(1)} + 1 \geq \tilde{k}$  and, by definition of  $\tilde{k}$ , for any  $j$  we have  $p_j \leq t_{\tilde{k}} \Leftrightarrow p_j \leq t_{\tilde{k}+1}$ . First prove that  $p_1 \leq t_{\tilde{k}} \Leftrightarrow \tilde{k} = \tilde{k}'_{(1)} + 1$ : using the definitions of  $\tilde{k}$  and  $\tilde{k}'_{(1)}$  we obtain  $p_1 \leq t_{\tilde{k}} \Leftrightarrow |\{2 \leq j \leq m \mid p_j \leq t_{\tilde{k}+1}\}| < \tilde{k} \Leftrightarrow \tilde{k}'_{(1)} < \tilde{k} \Leftrightarrow \tilde{k}'_{(1)} + 1 = \tilde{k}$ . Second, we prove  $p_1 > t_{\tilde{k}} \Leftrightarrow p_1 > t_{\tilde{k}(1)+1}$ : since we obviously have  $\tilde{k} \geq \tilde{k}_{(1)}$ , we get  $p_1 > t_{\tilde{k}} \Rightarrow p_1 > t_{\tilde{k}+1} \Rightarrow p_1 > t_{\tilde{k}(1)+1}$ . Conversely, if  $p_1 > t_{\tilde{k}(1)+1}$ , we get  $|\{j \mid p_j \leq t_{\tilde{k}(1)+1}\}| = |\{j \geq 2 \mid p_j \leq t_{\tilde{k}(1)+1}\}| < \tilde{k}_{(1)} + 1$  (by definition of  $\tilde{k}_{(1)}$ ), hence  $\tilde{k}_{(1)} + 1 > \tilde{k}$  (by definition of  $\tilde{k}$ ), which implies  $\tilde{k}_{(1)} = \tilde{k}$ , thus  $p_1 > t_{\tilde{k}+1}$  and finally  $p_1 > t_{\tilde{k}}$ .  $\square$

LEMMA A.3. For  $m \geq 1$ , let  $(X_1, Y_1), \dots, (X_m, Y_m)$  be exchangeable Bernoulli variable couples, i.e.  $((X_1, Y_1), \dots, (X_m, Y_m)) \sim ((X_{\sigma(1)}, Y_{\sigma(1)}), \dots, (X_{\sigma(m)}, Y_{\sigma(m)}))$  for any permutation  $\sigma$  of  $\{1, \dots, m\}$  and  $X_i, Y_i \in \{0, 1\}$ . Let us consider  $f : \{0, 1\}^m \rightarrow (0, \infty)$  a measurable function with exchangeable coordinates, that is, for any permutation  $\sigma$  of  $\{1, \dots, m\}$ , and any  $(x_1, \dots, x_m) \in \{0, 1\}^m$ , we have  $f(x_{\sigma(1)}, \dots, x_{\sigma(m)}) = f(x_1, \dots, x_m)$ . Then we have, for any  $s \geq 1$ ,

$$\mathbb{E} \left[ f(X_1, \dots, X_m)^s \left( \sum_{i=1}^m X_i Y_i \right)^s \right] = \sum_{\ell=1}^{s \wedge m} \frac{m!}{(m-\ell)!} \left\{ \begin{matrix} s \\ \ell \end{matrix} \right\} \mathbb{E}(f(X_1, \dots, X_m)^s X_1 Y_1 \dots X_\ell Y_\ell).$$

PROOF. We prove the result for  $s \leq m$  (the result for  $s > m$  is similar), using (11), by a recurrence on  $s$ : this is true for  $s = 1$ ; for the step  $s + 1$ , assume the property true for  $s$  and write

$$\mathbb{E} \left[ f(X_1, \dots, X_m)^{s+1} \left( \sum_{i=1}^m X_i Y_i \right)^{s+1} \right] = \sum_{j=1}^m \mathbb{E} \left[ (f(X_1, \dots, X_m)^{1+1/s})^s \left( \sum_{i=1}^m X_i X_j Y_i Y_j \right)^s \right],$$

so that we may apply the property at step  $s$  to variables  $(X_i X_j, Y_i Y_j)_i$ :

$$\begin{aligned}
& \mathbb{E} \left[ f(X_1, \dots, X_m)^{s+1} \left( \sum_{i=1}^m X_i Y_i \right)^{s+1} \right] \\
&= \sum_{j=1}^m \sum_{\ell=1}^s \frac{m!}{(m-\ell)!} \left\{ \begin{matrix} s \\ \ell \end{matrix} \right\} \mathbb{E}(f(X_1, \dots, X_m)^{s+1} X_1 Y_1 \dots X_\ell Y_\ell X_j Y_j) \\
&= \sum_{\ell=1}^s \sum_{j=1}^{\ell} \frac{m!}{(m-\ell)!} \left\{ \begin{matrix} s \\ \ell \end{matrix} \right\} \mathbb{E}(f(X_1, \dots, X_m)^{s+1} X_1 Y_1 \dots X_\ell Y_\ell) \\
&\quad + \sum_{\ell=1}^s \sum_{j=\ell+1}^m \frac{m!}{(m-\ell)!} \left\{ \begin{matrix} s \\ \ell \end{matrix} \right\} \mathbb{E}(f(X_1, \dots, X_m)^{s+1} X_1 Y_1 \dots X_\ell Y_\ell X_{\ell+1} Y_{\ell+1}) \\
&= \sum_{\ell=1}^s \frac{m!}{(m-\ell)!} \ell \left\{ \begin{matrix} s \\ \ell \end{matrix} \right\} \mathbb{E}(f(X_1, \dots, X_m)^{s+1} X_1 Y_1 \dots X_\ell Y_\ell) \\
&\quad + \sum_{\ell=1}^s \frac{m!}{(m-\ell-1)!} \left\{ \begin{matrix} s \\ \ell \end{matrix} \right\} \mathbb{E}(f(X_1, \dots, X_m)^{s+1} X_1 Y_1 \dots X_\ell Y_\ell X_{\ell+1} Y_{\ell+1}).
\end{aligned}$$

We conclude by using (11). □

UPMC UNIV PARIS 6, LPMA,  
4, PLACE JUSSIEU, 75252 PARIS CEDEX 05, FRANCE  
E-MAIL: etienne.roquain@upmc.fr

UPMC UNIV PARIS 6, LPMA,  
4, PLACE JUSSIEU, 75252 PARIS CEDEX 05, FRANCE  
E-MAIL: fanny.villers@upmc.fr